

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using fw model

Run on: January 31, 2003, 07:42:31 (Search time 73 Seconds
(without alignments))

1643.201 Million cell updates/sec

Title: US-09-807-201-7
Sequence: 1 gacatccagatccatcc.....cgacacccaatcccatcc 267

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 396772 seqs, 224632407 residues

Total number of hits satisfying chosen parameters: 793544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	267	100.0	683 9 US-10-012-896-896	Sequence 896, App
2	267	100.0	683 9 US-09-895-793-896	Sequence 896, App
3	267	100.0	683 9 US-09-895-814-896	Sequence 896, App
4	267	100.0	683 10 US-09-759-143-896	Sequence 896, App
5	267	100.0	683 10 US-09-780-669-896	Sequence 896, App
6	267	100.0	683 10 US-09-822-827-896	Sequence 896, App
7	267	100.0	2479 9 US-10-012-896-894	Sequence 894, App
8	267	100.0	2479 9 US-09-895-793-894	Sequence 894, App
9	267	100.0	2479 9 US-09-895-814-894	Sequence 894, App
10	267	100.0	2479 10 US-09-759-143-894	Sequence 894, App
11	267	100.0	2479 10 US-09-780-669-894	Sequence 894, App
12	267	100.0	2479 10 US-09-822-827-894	Sequence 894, App
13	267	100.0	3245 9 US-10-012-896-929	Sequence 929, App
14	267	100.0	3245 9 US-09-895-793-929	Sequence 929, App
15	267	100.0	3245 9 US-09-895-814-929	Sequence 929, App
16	267	100.0	3245 10 US-09-759-143-929	Sequence 929, App
17	267	100.0	3245 10 US-09-780-669-929	Sequence 929, App
18	267	100.0	3245 10 US-09-822-827-929	Sequence 929, App
19	265.4	99.4	2486 9 US-09-981-353-22	Sequence 22, App

20	219	82.0	1476 9 US-10-012-896-931	Sequence 931, App
21	219	82.0	1476 9 US-09-895-793-931	Sequence 931, App
22	219	82.0	1476 9 US-09-895-814-931	Sequence 931, App
23	219	82.0	1476 10 US-09-759-143-931	Sequence 931, App
24	219	82.0	1476 10 US-09-780-669-931	Sequence 931, App
25	219	82.0	1476 10 US-09-822-827-931	Sequence 931, App
26	219	82.0	1479 9 US-10-012-896-930	Sequence 930, App
27	219	82.0	1479 9 US-09-895-793-930	Sequence 930, App
28	219	82.0	1479 9 US-09-895-814-930	Sequence 930, App
29	219	82.0	1479 10 US-09-759-143-930	Sequence 930, App
30	219	82.0	1479 10 US-09-780-669-930	Sequence 930, App
31	219	82.0	1479 10 US-09-822-827-930	Sequence 930, App
32	178	66.7	196 9 US-09-232-880-151	Sequence 151, App
33	178	66.7	196 9 US-10-012-896-151	Sequence 151, App
34	178	66.7	196 9 US-09-895-793-151	Sequence 151, App
35	178	66.7	196 9 US-09-895-814-151	Sequence 151, App
36	178	66.7	196 10 US-09-759-143-151	Sequence 151, App
37	178	66.7	196 10 US-09-780-669-151	Sequence 151, App
38	178	66.7	196 10 US-09-822-827-151	Sequence 151, App
39	178	66.7	196 10 US-09-822-827-151	Sequence 151, App
40	178	66.7	196 10 US-09-115-453-151	Sequence 151, App
41	32.2	12.1	369 10 US-09-860-352-4371	Sequence 4371, App
42	32.2	12.1	412 10 US-09-864-761-22443	Sequence 22443, App
43	32.2	12.1	442 10 US-09-864-761-5680	Sequence 5680, App
44	32	12.0	2479 9 US-10-012-896-894	Sequence 894, App
45	32	12.0	2479 9 US-09-895-793-894	Sequence 894, App

ALIGNMENTS

RESULT 1
US-10-012-896-896
Sequence 896, Application US/10012896
Publication No. US20020183251A1
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan L.
APPLICANT: Jiang, Yugu
APPLICANT: Kalos, Michael D.
APPLICANT: Retter, Marc W.
APPLICANT: Stolk, John A.
APPLICANT: Day, Craig H.
APPLICANT: Vedvick, Thomas S.
APPLICANT: Carter, Darrick
APPLICANT: Li, Samuel X.
APPLICANT: Wang, Aijun
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Hepler, William T.
APPLICANT: Henderson, Robert A.
APPLICANT: Hurral, John
APPLICANT: McNeill, Patricia D.
APPLICANT: Houghton, Raymond L.
APPLICANT: Vinals de Bassols, Carlota
APPLICANT: Foy, Teresa
APPLICANT: Foy, Gary R.
APPLICANT: Mantabe, Yoshitiro
APPLICANT: Meagher, Madeleine Joy
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.427C27
CURRENT APPLICATION NUMBER: US/10/012, 896
CURRENT FILING DATE: 2001-12-10
NUMBER OF SEQ ID NOS: 1011
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 896
LENGTH: 683
TYPE: DNA
ORGANISM: Homo sapiens
US-10-012-896-896

Query Match 100.0%; Score 267; DB 9; Length 683;
 Best Local Similarity 100.0%; Pred. No. 5.7e-75;
 Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAACATTCAGATACCTATCTATCTGATGCTGTGATAACAGCAAGATGGCTTTGAAC 60
 DB 9 GAACATTCAGATACCTATCTATCTGATGCTGTGATAACAGCAAGATGGCTTTGAAC 68
 QY 61 TCAGGGTCACACAGCTATTTGACCTTACTATGAAAACCATGATACCAACCGGAAAAC 120
 DB 69 TCAGGGTCACACAGCTATTTGACCTTACTATGAAAACCATGATACCAACCGGAAAAC 128
 QY 121 CCTATTCGGCAGACGCTGTGTGCTCCCACTGTCTACAGAGTGATCCGGCTCAGTAC 180
 DB 129 CCTATTCGGCAGACGCTGTGTGCTCCCACTGTCTACAGAGTGATCCGGCTCAGTAC 188
 QY 181 TACCCGTCCCGGTGCGCCAGTACGCGCCGAGGGTCTCTGACGAGGTTCCACCCCGTC 240
 DB 189 TACCCGTCCCGGTGCGCCAGTACGCGCCGAGGGTCTCTGACGAGGTTCCACCCCGTC 248
 QY 241 GTCTGACGACGACCCCAATCCCATCC 267
 DB 249 GTCTGACGACGACCCCAATCCCATCC 275

RESULT 2

US-09-895-793-896
 Sequence 896, Application US/09895793
 Publication No. US20020192763A1

GENERAL INFORMATION:

APPLICANT: Xu, Jiangchun
 APPLICANT: Dillon, Davin C.
 APPLICANT: Mitcham, Jennifer L.
 APPLICANT: Harlocker, Susan L.
 APPLICANT: Jiang, Yudi
 APPLICANT: Kalos, Michael D.
 APPLICANT: Retter, Marc W.
 APPLICANT: Stoik, John A.
 APPLICANT: Day, Craig H.
 APPLICANT: Vedvick, Thomas S.
 APPLICANT: Carter, Darrick
 APPLICANT: Li, Samuel X.
 APPLICANT: Wang, Aijun
 APPLICANT: Skeiky, Yasir A.W.
 APPLICANT: Hepler, William T.
 APPLICANT: Henderson, Robert A.
 APPLICANT: Hural, John
 APPLICANT: McNeill, Patricia D.
 APPLICANT: Houghton, Raymond L.
 APPLICANT: Vinals de Basols, Carlota
 APPLICANT: Foy, Teresa
 APPLICANT: Fanger, Gary R.
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
 TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
 FILE REFERENCE: 210121.534C2
 CURRENT APPLICATION NUMBER: US/09/895,793
 CURRENT FILING DATE: 2001-06-29
 NUMBER OF SEQ ID NOS: 982
 SOFTWARE: FastSeq for Windows Version 3.0
 SEQ ID NO 896
 LENGTH: 683
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-09-895-793-896

Query Match 100.0%; Score 267; DB 9; Length 683;
 Best Local Similarity 100.0%; Pred. No. 5.7e-75;
 Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAACATTCAGATACCTATCTATCTGATGCTGTGATAACAGCAAGATGGCTTTGAAC 60
 DB 9 GAACATTCAGATACCTATCTATCTGATGCTGTGATAACAGCAAGATGGCTTTGAAC 68

QY 61 TCAGGGTCACACAGCTATTTGACCTTACTATGAAAACCATGATACCAACCGGAAAAC 120
 DB 69 TCAGGGTCACACAGCTATTTGACCTTACTATGAAAACCATGATACCAACCGGAAAAC 128
 QY 121 CCTATTCGGCAGACGCTGTGTGCTCCCACTGTCTACAGAGTGATCCGGCTCAGTAC 180
 DB 129 CCTATTCGGCAGACGCTGTGTGCTCCCACTGTCTACAGAGTGATCCGGCTCAGTAC 188
 QY 181 TACCCGTCCCGGTGCGCCAGTACGCGCCGAGGGTCTCTGACGAGGTTCCACCCCGTC 240
 DB 189 TACCCGTCCCGGTGCGCCAGTACGCGCCGAGGGTCTCTGACGAGGTTCCACCCCGTC 248
 QY 241 GTCTGACGACGACCCCAATCCCATCC 267
 DB 249 GTCTGACGACGACCCCAATCCCATCC 275

RESULT 3

US-09-895-814-896
 Sequence 896, Application US/09895814
 Publication No. US20020193296A1

GENERAL INFORMATION:

APPLICANT: Xu, Jiangchun
 APPLICANT: Dillon, Davin C.
 APPLICANT: Mitcham, Jennifer L.
 APPLICANT: Harlocker, Susan L.
 APPLICANT: Jiang, Yudi
 APPLICANT: Kalos, Michael D.
 APPLICANT: Retter, Marc W.
 APPLICANT: Stoik, John A.
 APPLICANT: Day, Craig H.
 APPLICANT: Vedvick, Thomas S.
 APPLICANT: Carter, Darrick
 APPLICANT: Li, Samuel X.
 APPLICANT: Wang, Aijun
 APPLICANT: Skeiky, Yasir A.W.
 APPLICANT: Hepler, William T.
 APPLICANT: Henderson, Robert A.
 APPLICANT: Hural, John
 APPLICANT: McNeill, Patricia D.
 APPLICANT: Houghton, Raymond L.
 APPLICANT: Vinals de Basols, Carlota
 APPLICANT: Foy, Teresa
 APPLICANT: Fanger, Gary R.
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
 TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
 FILE REFERENCE: 210121.427C26
 CURRENT APPLICATION NUMBER: US/09/895,814
 CURRENT FILING DATE: 2001-06-29
 NUMBER OF SEQ ID NOS: 990
 SOFTWARE: FastSeq for Windows Version 3.0
 SEQ ID NO 896
 LENGTH: 683
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-09-895-814-896

Query Match 100.0%; Score 267; DB 9; Length 683;
 Best Local Similarity 100.0%; Pred. No. 5.7e-75;
 Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAACATTCAGATACCTATCTATCTGATGCTGTGATAACAGCAAGATGGCTTTGAAC 60
 DB 9 GAACATTCAGATACCTATCTATCTGATGCTGTGATAACAGCAAGATGGCTTTGAAC 68
 QY 61 TCAGGGTCACACAGCTATTTGACCTTACTATGAAAACCATGATACCAACCGGAAAAC 120
 DB 69 TCAGGGTCACACAGCTATTTGACCTTACTATGAAAACCATGATACCAACCGGAAAAC 128
 QY 121 CCTATTCGGCAGACGCTGTGTGCTCCCACTGTCTACAGAGTGATCCGGCTCAGTAC 180
 DB 129 CCTATTCGGCAGACGCTGTGTGCTCCCACTGTCTACAGAGTGATCCGGCTCAGTAC 188

QY 181 TACCGTCCCCCGTGGCCGAGTACGCCCCGAGGGTCTGACGAGGCTTCCAAACCCCGTC 240
 DB 189 TACCGTCCCCCGTGGCCGAGTACGCCCCGAGGGTCTGACGAGGCTTCCAAACCCCGTC 248
 QY 241 GTCTGACGACGAGCCCAATCCCATCC 267
 DB 249 GTCTGACGACGAGCCCAATCCCATCC 275

RESULT 4

US-09-759-143-896
 ; Sequence 896, Application US/09759143
 ; Patent No. US2002002248A1
 ; GENERAL INFORMATION:

APPLICANT: Xu, Jiangchun
 APPLICANT: Dillon, Devin C.
 APPLICANT: Mitcham, Jennifer L.
 APPLICANT: Harlocker, Susan L.
 APPLICANT: Jiang, Yugu
 APPLICANT: Henderson, Robert A.
 APPLICANT: Kalos, Michael D.
 APPLICANT: Fanger, Gary R.
 APPLICANT: Retter, Marc W.
 APPLICANT: Stolk, John A.
 APPLICANT: Day, Craig H.
 APPLICANT: Vedvick, Thomas S.
 APPLICANT: Carter, Darick
 APPLICANT: Li, Samuel
 APPLICANT: Wang, Aijun
 APPLICANT: Skeiky, Yasir A.W.
 APPLICANT: Hepler, William
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
 FILE REFERENCE: 210121.427C23
 CURRENT APPLICATION NUMBER: US/09/759,143
 CURRENT FILING DATE: 2001-01-12
 NUMBER OF SEQ ID NOS: 934
 SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 896
 ; LENGTH: 683
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-759-143-896

Query Match 100.0%; Score 267; DB 10; Length 683;
 Best Local Similarity 100.0%; Pred. No. 5,7e-75;
 Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAACATTCAGATACCTTACTCATCTGATGCTGTTGATTAACGACGAATGGCTTTGAAC 60
 DB 9 GAACATTCAGATACCTTACTCATCTGATGCTGTTGATTAACGACGAATGGCTTTGAAC 68
 QY 61 TCAGGGTACACCAAGCTATTGAGCTTACTATGAAAAACAATGATACCAACCGGAAAAAC 120
 DB 69 TCAGGGTACACCAAGCTATTGAGCTTACTATGAAAAACAATGATACCAACCGGAAAAAC 128
 QY 121 CCCTATCCCGACAGCCCACTGTGTCTCCCACTGTCTAAGAGTGCATCCGGCTCAGTAC 180
 DB 129 CCCTATCCCGACAGCCCACTGTGTCTCCCACTGTCTAAGAGTGCATCCGGCTCAGTAC 188
 QY 181 TACCGTCCCCCGTGGCCGAGTACGCCCCGAGGGTCTGACGAGGCTTCCAAACCCCGTC 240
 DB 189 TACCGTCCCCCGTGGCCGAGTACGCCCCGAGGGTCTGACGAGGCTTCCAAACCCCGTC 248
 QY 241 GTCTGACGACGAGCCCAATCCCATCC 267
 DB 249 GTCTGACGACGAGCCCAATCCCATCC 275

RESULT 5

US-09-780-669-896
 ; Sequence 896, Application US/09780669
 ; Patent No. US2002005197A1

GENERAL INFORMATION:

APPLICANT: Xu, Jiangchun
 APPLICANT: Dillon, Devin C.
 APPLICANT: Mitcham, Jennifer L.
 APPLICANT: Harlocker, Susan L.
 APPLICANT: Jiang, Yugu
 APPLICANT: Henderson, Robert A.
 APPLICANT: Kalos, Michael D.
 APPLICANT: Fanger, Gary R.
 APPLICANT: Retter, Marc W.
 APPLICANT: Stolk, John A.
 APPLICANT: Day, Craig H.
 APPLICANT: Vedvick, Thomas S.
 APPLICANT: Carter, Darick
 APPLICANT: Li, Samuel
 APPLICANT: Wang, Aijun
 APPLICANT: Skeiky, Yasir A.W.
 APPLICANT: Hepler, William
 APPLICANT: Hurrell, John
 APPLICANT: McNeill, Patricia D.
 APPLICANT: Houghton, Raymond L.
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
 FILE REFERENCE: 210121.427C24
 CURRENT APPLICATION NUMBER: US/09/780,669
 CURRENT FILING DATE: 2001-02-09
 NUMBER OF SEQ ID NOS: 943
 SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 896
 ; LENGTH: 683
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-780-669-896

Query Match 100.0%; Score 267; DB 10; Length 683;
 Best Local Similarity 100.0%; Pred. No. 5,7e-75;
 Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAACATTCAGATACCTTACTCATCTGATGCTGTTGATTAACGACGAATGGCTTTGAAC 60
 DB 9 GAACATTCAGATACCTTACTCATCTGATGCTGTTGATTAACGACGAATGGCTTTGAAC 68
 QY 61 TCAGGGTACACCAAGCTATTGAGCTTACTATGAAAAACAATGATACCAACCGGAAAAAC 120
 DB 69 TCAGGGTACACCAAGCTATTGAGCTTACTATGAAAAACAATGATACCAACCGGAAAAAC 128
 QY 121 CCCTATCCCGACAGCCCACTGTGTCTCCCACTGTCTAAGAGTGCATCCGGCTCAGTAC 180
 DB 129 CCCTATCCCGACAGCCCACTGTGTCTCCCACTGTCTAAGAGTGCATCCGGCTCAGTAC 188
 QY 181 TACCGTCCCCCGTGGCCGAGTACGCCCCGAGGGTCTGACGAGGCTTCCAAACCCCGTC 240
 DB 189 TACCGTCCCCCGTGGCCGAGTACGCCCCGAGGGTCTGACGAGGCTTCCAAACCCCGTC 248
 QY 241 GTCTGACGACGAGCCCAATCCCATCC 267
 DB 249 GTCTGACGACGAGCCCAATCCCATCC 275

RESULT 6

US-09-822-827-896
 ; Sequence 896, Application US/09822827
 ; Patent No. US20020081680A1
 ; GENERAL INFORMATION:

APPLICANT: Xu, Jiangchun
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
 FILE REFERENCE: 210121.534C1
 CURRENT APPLICATION NUMBER: US/09/822,827
 CURRENT FILING DATE: 2001-03-28
 NUMBER OF SEQ ID NOS: 982
 SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 896

LENGTH: 683
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-09-822-827-896

Query Match 100.0%; Score 267; DB 10; Length 683;
 Best Local Similarity 100.0%; Pred. No. 5,7e-75;
 Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAACATTCAGATACCTATCATCTACCTGCTGTGATACAGCAAGATGGCTTGAAC 60
 DB 9 GAACATTCAGATACCTATCATCTACCTGCTGTGATACAGCAAGATGGCTTGAAC 68
 QY 61 TCAGGATCAGCAGCAGCTATTTGAGCCTTACTATGAAGACATGATACCAACCGGAAAC 120
 DB 69 TCAGGATCAGCAGCAGCTATTTGAGCCTTACTATGAAGACATGATACCAACCGGAAAC 128
 QY 121 CCTATCCGACAGCCCACTGTGTGCCCACTGTCTACAGAGTGATCCGGCTCAGTAC 180
 DB 129 CCTATCCGACAGCCCACTGTGTGCCCACTGTCTACAGAGTGATCCGGCTCAGTAC 188
 QY 181 TACCGGTCCCGGTCCGAGTACGCGGAGGCTCTGACAGAGGCTTCCAAACCCCGTC 240
 DB 189 TACCGGTCCCGGTCCGAGTACGCGGAGGCTCTGACAGAGGCTTCCAAACCCCGTC 248
 QY 241 GTCTGACAGCAGCCCAATCCCATCC 267
 DB 249 GTCTGACAGCAGCCCAATCCCATCC 275

RESULT 7

US-10-012-896-894
 Sequence 894, Application US/10012896
 Publication No. US20020183251A1

GENERAL INFORMATION:

APPLICANT: Xu, Jiangchun
 APPLICANT: Dillon, David C.
 APPLICANT: Mitcham, Jennifer L.
 APPLICANT: Harlocker, Susan L.
 APPLICANT: Jiang, Yugu
 APPLICANT: Kalos, Michael D.
 APPLICANT: Retter, Marc W.
 APPLICANT: Stolk, John A.
 APPLICANT: Day, Craig H.
 APPLICANT: Vedvick, Thomas S.
 APPLICANT: Carter, Derrick
 APPLICANT: Li, Samuel X.
 APPLICANT: Wang, Aijun
 APPLICANT: Skeiky, Yasir A.W.
 APPLICANT: Hepler, William T.
 APPLICANT: Henderson, Robert A.
 APPLICANT: Hural, John
 APPLICANT: McNeill, Patricia D.
 APPLICANT: Houghton, Raymond L.
 APPLICANT: Vinals de Bassols, Carlota
 APPLICANT: Foy, Teresa
 APPLICANT: Fanger, Gary R.
 APPLICANT: Mantanabe, Yoshihiro
 APPLICANT: Meagher, Madeleine Joy
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
 TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
 FILE REFERENCE: 210121.427C27
 CURRENT APPLICATION NUMBER: US/10/012,896
 CURRENT FILING DATE: 2001-12-10
 NUMBER OF SEQ ID NOS: 1011
 SOFTWARE: FastSeq for Windows Version 3.0
 SEQ ID NO 894
 LENGTH: 2479
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-10-012-896-894

Query Match 100.0%; Score 267; DB 9; Length 2479;

Best Local Similarity 100.0%; Pred. No. 9,4e-75;
 Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAACATTCAGATACCTATCATCTACCTGCTGTGATACAGCAAGATGGCTTGAAC 60
 DB 9 GAACATTCAGATACCTATCATCTACCTGCTGTGATACAGCAAGATGGCTTGAAC 68
 QY 61 TCAGGATCAGCAGCAGCTATTTGAGCCTTACTATGAAGACATGATACCAACCGGAAAC 120
 DB 69 TCAGGATCAGCAGCAGCTATTTGAGCCTTACTATGAAGACATGATACCAACCGGAAAC 128
 QY 121 CCTATCCGACAGCCCACTGTGTGCCCACTGTCTACAGAGTGATCCGGCTCAGTAC 180
 DB 129 CCTATCCGACAGCCCACTGTGTGCCCACTGTCTACAGAGTGATCCGGCTCAGTAC 188
 QY 181 TACCGGTCCCGGTCCGAGTACGCGGAGGCTCTGACAGAGGCTTCCAAACCCCGTC 240
 DB 189 TACCGGTCCCGGTCCGAGTACGCGGAGGCTCTGACAGAGGCTTCCAAACCCCGTC 248
 QY 241 GTCTGACAGCAGCCCAATCCCATCC 267
 DB 249 GTCTGACAGCAGCCCAATCCCATCC 275

RESULT 8

US-09-895-793-894
 Sequence 894, Application US/09895793
 Publication No. US20020192763A1

GENERAL INFORMATION:

APPLICANT: Xu, Jiangchun
 APPLICANT: Dillon, David C.
 APPLICANT: Mitcham, Jennifer L.
 APPLICANT: Harlocker, Susan L.
 APPLICANT: Jiang, Yugu
 APPLICANT: Kalos, Michael D.
 APPLICANT: Retter, Marc W.
 APPLICANT: Stolk, John A.
 APPLICANT: Day, Craig H.
 APPLICANT: Vedvick, Thomas S.
 APPLICANT: Carter, Derrick
 APPLICANT: Li, Samuel X.
 APPLICANT: Wang, Aijun
 APPLICANT: Skeiky, Yasir A.W.
 APPLICANT: Hepler, William T.
 APPLICANT: Henderson, Robert A.
 APPLICANT: Hural, John
 APPLICANT: McNeill, Patricia D.
 APPLICANT: Houghton, Raymond L.
 APPLICANT: Vinals de Bassols, Carlota
 APPLICANT: Foy, Teresa
 APPLICANT: Fanger, Gary R.
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
 TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
 FILE REFERENCE: 210121.534C2
 CURRENT APPLICATION NUMBER: US/09/895,793
 CURRENT FILING DATE: 2001-06-29
 NUMBER OF SEQ ID NOS: 982
 SOFTWARE: FastSeq for Windows Version 3.0
 SEQ ID NO 894
 LENGTH: 2479
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-09-895-793-894

Query Match 100.0%; Score 267; DB 9; Length 2479;
 Best Local Similarity 100.0%; Pred. No. 9,4e-75;
 Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAACATTCAGATACCTATCATCTACCTGCTGTGATACAGCAAGATGGCTTGAAC 60
 DB 9 GAACATTCAGATACCTATCATCTACCTGCTGTGATACAGCAAGATGGCTTGAAC 68
 QY 61 TCAGGATCAGCAGCAGCTATTTGAGCCTTACTATGAAGACATGATACCAACCGGAAAC 120

Db 69 TCAGGATACACCAAGTATGACCTTACTATGAAAACCAAGATACCAACCGGAAAAC 128
OY 121 CCTATCCCGACACCCCACTGTGTCTCCCATCTGTCTACAGAGTGCATCCGGCTCAATAC 180
Db 129 CCTATCCCGACACCCCACTGTGTCTCCCATCTGTCTACAGAGTGCATCCGGCTCAATAC 188
OY 181 TACCCGTCCCGGTGCGCCGATACGCGCCGAGAGGTCTGTGACGACAGGCTTCCAAACCCGCTC 240
Db 189 TACCCGTCCCGGTGCGCCGATACGCGCCGAGAGGTCTGTGACGACAGGCTTCCAAACCCGCTC 248
OY 241 GTCTGACGACGACCCCAATCCCATCC 267
Db 249 GTCTGACGACGACCCCAATCCCATCC 275

RESULT 9
US-09-895-814-894
Sequence 894, Application US/09895814
Publication No. US20020193296A1
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan L.
APPLICANT: Jiang, Yugu
APPLICANT: Kalos, Michael D.
APPLICANT: Retter, Marc W.
APPLICANT: Stolck, John A.
APPLICANT: Day, Craig H.
APPLICANT: Vedvick, Thomas S.
APPLICANT: Carter, Darick
APPLICANT: Li, Samuel X.
APPLICANT: Wang, Aijun
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Hepler, William T.
APPLICANT: Henderson, Robert A.
APPLICANT: Hurai, John
APPLICANT: McNeill, Patricia D.
APPLICANT: Houghton, Raymond L.
APPLICANT: Vinals de Bassols, Carlota
APPLICANT: Foy, Teresa
APPLICANT: Fanger, Gary R.
APPLICANT: Fanger, Gary R.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
FILE REFERENCE: 210121.427C26
CURRENT APPLICATION NUMBER: US/09/895,814
CURRENT FILING DATE: 2001-06-29
NUMBER OF SEQ ID NOS: 990
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 894
LENGTH: 2479
TYPE: DNA
ORGANISM: Homo sapiens
US-09-895-814-894

Query Match 100.0%; Score 267; DB 9; Length 2479;
Best Local Similarity 100.0%; Pred. No. 9.4e-75;
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAACATTCGATACCTATGATCTGATGCTGTGATAAACAAGATGCTTTGAAC 60
Db 9 GAACATTCGATACCTATGATCTGATGCTGTGATAAACAAGATGCTTTGAAC 68
OY 61 TCAGGATACACCAAGTATGACCTTACTATGAAAACCAAGATACCAACCGGAAAAC 120
Db 69 TCAGGATACACCAAGTATGACCTTACTATGAAAACCAAGATACCAACCGGAAAAC 128
OY 121 CCTATCCCGACACCCCACTGTGTCTCCCATCTGTCTACAGAGTGCATCCGGCTCAATAC 180
Db 129 CCTATCCCGACACCCCACTGTGTCTCCCATCTGTCTACAGAGTGCATCCGGCTCAATAC 188
OY 181 TACCCGTCCCGGTGCGCCGATACGCGCCGAGAGGTCTGTGACGACAGGCTTCCAAACCCGCTC 240
Db 189 TACCCGTCCCGGTGCGCCGATACGCGCCGAGAGGTCTGTGACGACAGGCTTCCAAACCCGCTC 248
OY 241 GTCTGACGACGACCCCAATCCCATCC 267
Db 249 GTCTGACGACGACCCCAATCCCATCC 275

Db 189 TACCCGTCCCGGTGCGCCGATACGCGCCGAGAGGTCTGTGACGACAGGCTTCCAAACCCGCTC 248
OY 241 GTCTGACGACGACCCCAATCCCATCC 267
Db 249 GTCTGACGACGACCCCAATCCCATCC 275

RESULT 10
US-09-759-143-894
Sequence 894, Application US/09759143
Patent No. US20020022248A1
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan L.
APPLICANT: Jiang, Yugu
APPLICANT: Henderson, Robert A.
APPLICANT: Kalos, Michael D.
APPLICANT: Retter, Marc W.
APPLICANT: Stolck, John A.
APPLICANT: Day, Craig H.
APPLICANT: Vedvick, Thomas S.
APPLICANT: Carter, Darick
APPLICANT: Li, Samuel
APPLICANT: Wang, Aijun
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Hepler, William
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
FILE REFERENCE: 210121.427C23
CURRENT APPLICATION NUMBER: US/09/759,143
CURRENT FILING DATE: 2001-01-12
NUMBER OF SEQ ID NOS: 934
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 894
LENGTH: 2479
TYPE: DNA
ORGANISM: Homo sapiens
US-09-759-143-894

Query Match 100.0%; Score 267; DB 10; Length 2479;
Best Local Similarity 100.0%; Pred. No. 9.4e-75;
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAACATTCGATACCTATGATCTGATGCTGTGATAAACAAGATGCTTTGAAC 60
Db 9 GAACATTCGATACCTATGATCTGATGCTGTGATAAACAAGATGCTTTGAAC 68
OY 61 TCAGGATACACCAAGTATGACCTTACTATGAAAACCAAGATACCAACCGGAAAAC 120
Db 69 TCAGGATACACCAAGTATGACCTTACTATGAAAACCAAGATACCAACCGGAAAAC 128
OY 121 CCTATCCCGACACCCCACTGTGTCTCCCATCTGTCTACAGAGTGCATCCGGCTCAATAC 180
Db 129 CCTATCCCGACACCCCACTGTGTCTCCCATCTGTCTACAGAGTGCATCCGGCTCAATAC 188
OY 181 TACCCGTCCCGGTGCGCCGATACGCGCCGAGAGGTCTGTGACGACAGGCTTCCAAACCCGCTC 240
Db 189 TACCCGTCCCGGTGCGCCGATACGCGCCGAGAGGTCTGTGACGACAGGCTTCCAAACCCGCTC 248
OY 241 GTCTGACGACGACCCCAATCCCATCC 267
Db 249 GTCTGACGACGACCCCAATCCCATCC 275

RESULT 11
US-09-780-669-894
Sequence 894, Application US/09780669
Patent No. US20020051977A1
GENERAL INFORMATION:

```

: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-822-827--894

Query Match      100.0%; Score 267; DB 10; Length 2479;
Best Local Similarity 100.0%; Pident. No. 9,4e-75;
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  GAACATTCAGATTCCTTCACTTACTGCATGCTGTGTAATPACAGCAAGATGGCTTTGAAC 60
DB      9  GAACTATTCAGATCTTATCACTTACTGCATGCTGTGTAATPACAGCAAGATGGCTTTGAAC 68
QY      61  TCAGGTCACCAACCAAGCTATTGACCTTACTATGAAAAACATGATATCAACACCGAAAAAC 120
DB      69  TCAGGTCACCAACCAAGCTATTGACCTTACTATGAAAAACATGATATCAACACCGAAAAAC 128
QY      121  CCCATATCCCGGACAGCCGCACTGTGGTCCCGCACTGTCTCTGAGAGTGCATCCGGCTAGTAC 180
DB      129  CCCATATCCCGGACAGCCGCACTGTGGTCCCGCACTGTCTCTGAGAGTGCATCCGGCTAGTAC 188
QY      181  TACCCGTCCTCCCGTGGTCCCGCAAGTACGCGCCCGAGAGGCTCTGACGCGAGCTTCCAAACCCCGTC 240
DB      189  TACCCGTCCTCCCGTGGTCCCGCAAGTACCGCCCGGAGGGCTCTGACGCGAGCTTCCAAACCCCGTC 248
QY      241  GTCTGACGACGACGCCCAATCCCATCC 267
DB      249  GTCTGACGACGACGCCCAATCCCATCC 275

RESULT 13
US-10-012-896--929
: Sequence 929; Application US/10012896

```

```

GENERAL INFORMATION:
APPLICANT: Xu, Jianshun
APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan L.
APPLICANT: Jiang, Yugu
APPLICANT: Kalos, Michael D.
APPLICANT: Retter, Marc W.
APPLICANT: Stolk, John A.
APPLICANT: Day, Craig H.
APPLICANT: Vedvick, Thomas S.
APPLICANT: Carter, Darrick
APPLICANT: Li, Samuel X.
APPLICANT: Wang, Aijun
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Hepler, William T.
APPLICANT: Henderson, Robert A.
APPLICANT: Hural, John
APPLICANT: McNeill, Patricia D.
APPLICANT: Houghton, Raymond L.
APPLICANT: Vinals de Basbols, Carlota
APPLICANT: Foy, Teresa
APPLICANT: Fanger, Gary R.
APPLICANT: Wantanabe, Yoshihiro
APPLICANT: Meagher, Madeline Joy
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.427C27
CURRENT APPLICATION NUMBER: US/10/012,896
CURRENT FILING DATE: 2001-12-10
NUMBER OF SEQ. ID NOS: 101
SOFTWARE: fastseq for Windows Version 3.0
SEQ. ID NO 929
LENGTH: 3245
TYPE: DNA
ORGANISM: Homo sapiens
US-10-012-896-929

Query Match          100.0%; Score 267; DB 9; Length 3245;
Best Local Similarity 100.0%; Pred No. 1e-74;

```

Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAACATTCGAGATACCTATCTATCTGATGCTGTTGATAACAGAGATGCTTTGAAC 60
DB 97 GAACATTCGAGATACCTATCTATCTGATGCTGTTGATAACAGAGATGCTTTGAAC 156

QY 61 TCAGGGTCCACACAGCTATTTGACCTTATCTATGAAAAACATGATACCAACCGGAAAC 120
DB 157 TCAGGGTCCACACAGCTATTTGACCTTATCTATGAAAAACATGATACCAACCGGAAAC 216

QY 121 CCTATATCCGACAGACCCCACTGTGCTCCCACTGTCTACGAGTGCAATCCGGCTCAGTAC 180
DB 217 CCTATATCCGACAGACCCCACTGTGCTCCCACTGTCTACGAGTGCAATCCGGCTCAGTAC 276

QY 181 TACCCGTCCTCCCGTCCGACGAGTCCCGGAGGTCCTGACGAGGCTTCCAAACCCCGTC 240
DB 277 TACCCGTCCTCCCGTCCGACGAGTCCCGGAGGTCCTGACGAGGCTTCCAAACCCCGTC 336

QY 241 GTCGCAAGCAGACCCCAATCCCATCC 267
DB 337 GTCGCAAGCAGACCCCAATCCCATCC 363

RESULT 14

US-09-895-793-929
Sequence 929, Application US/09895793
Publication No. US20020192763A1
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan L.
APPLICANT: Jiang, Yugu
APPLICANT: Kalos, Michael D.
APPLICANT: Retter, Marc W.
APPLICANT: Stolk, John A.
APPLICANT: Day, Craig H.
APPLICANT: Vedvick, Thomas S.
APPLICANT: Carter, Darrick
APPLICANT: Li, Samuel X.
APPLICANT: Wang, Aijun
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Hepler, William T.
APPLICANT: Henderson, Robert A.
APPLICANT: Hural, John
APPLICANT: McNeill, Patricia D.
APPLICANT: Houghton, Raymond L.
APPLICANT: Vinals de Basols, Carlota
APPLICANT: Foy, Teresa
APPLICANT: Fanger, Gary R.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
FILE REFERENCE: 210121.534C2
CURRENT APPLICATION NUMBER: US/09/895,793
CURRENT FILING DATE: 2001-06-29
NUMBER OF SEQ ID NOS: 982
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 929
LENGTH: 3245
TYPE: DNA
ORGANISM: Homo sapiens
US-09-895-793-929

Query Match 100.0%; Score 267; DB 9; Length 3245;
Best Local Similarity 100.0%; Pred. No. 1e-74; Indels 0; Gaps 0;

Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAACATTCGAGATACCTATCTATCTGATGCTGTTGATAACAGAGATGCTTTGAAC 60
DB 97 GAACATTCGAGATACCTATCTATCTGATGCTGTTGATAACAGAGATGCTTTGAAC 156

QY 61 TCAGGGTCCACACAGCTATTTGACCTTATCTATGAAAAACATGATACCAACCGGAAAC 120
DB 157 TCAGGGTCCACACAGCTATTTGACCTTATCTATGAAAAACATGATACCAACCGGAAAC 216

DB 157 TCAGGGTCCACACAGCTATTTGACCTTATCTATGAAAAACATGATACCAACCGGAAAC 216
QY 121 CCTATATCCGACAGACCCCACTGTGCTCCCACTGTCTACGAGTGCAATCCGGCTCAGTAC 180
DB 217 CCTATATCCGACAGACCCCACTGTGCTCCCACTGTCTACGAGTGCAATCCGGCTCAGTAC 276

QY 181 TACCCGTCCTCCCGTCCGACGAGTCCCGGAGGTCCTGACGAGGCTTCCAAACCCCGTC 240
DB 277 TACCCGTCCTCCCGTCCGACGAGTCCCGGAGGTCCTGACGAGGCTTCCAAACCCCGTC 336

QY 241 GTCGCAAGCAGACCCCAATCCCATCC 267
DB 337 GTCGCAAGCAGACCCCAATCCCATCC 363

RESULT 15

US-09-895-814-929
Sequence 929, Application US/09895814
Publication No. US20020193236A1
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan L.
APPLICANT: Jiang, Yugu
APPLICANT: Kalos, Michael D.
APPLICANT: Retter, Marc W.
APPLICANT: Stolk, John A.
APPLICANT: Day, Craig H.
APPLICANT: Vedvick, Thomas S.
APPLICANT: Carter, Darrick
APPLICANT: Li, Samuel X.
APPLICANT: Wang, Aijun
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Hepler, William T.
APPLICANT: Henderson, Robert A.
APPLICANT: Hural, John
APPLICANT: McNeill, Patricia D.
APPLICANT: Houghton, Raymond L.
APPLICANT: Vinals de Basols, Carlota
APPLICANT: Foy, Teresa
APPLICANT: Fanger, Gary R.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
FILE REFERENCE: 210121.427C26
CURRENT APPLICATION NUMBER: US/09/895,814
CURRENT FILING DATE: 2001-06-29
NUMBER OF SEQ ID NOS: 990
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 929
LENGTH: 3245
TYPE: DNA
ORGANISM: Homo sapiens
US-09-895-814-929

Query Match 100.0%; Score 267; DB 9; Length 3245;
Best Local Similarity 100.0%; Pred. No. 1e-74; Indels 0; Gaps 0;

Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAACATTCGAGATACCTATCTATCTGATGCTGTTGATAACAGAGATGCTTTGAAC 60
DB 97 GAACATTCGAGATACCTATCTATCTGATGCTGTTGATAACAGAGATGCTTTGAAC 156

QY 61 TCAGGGTCCACACAGCTATTTGACCTTATCTATGAAAAACATGATACCAACCGGAAAC 120
DB 157 TCAGGGTCCACACAGCTATTTGACCTTATCTATGAAAAACATGATACCAACCGGAAAC 216

QY 121 CCTATATCCGACAGACCCCACTGTGCTCCCACTGTCTACGAGTGCAATCCGGCTCAGTAC 180
DB 217 CCTATATCCGACAGACCCCACTGTGCTCCCACTGTCTACGAGTGCAATCCGGCTCAGTAC 276

QY 181 TACCCGTCCTCCCGTCCGACGAGTCCCGGAGGTCCTGACGAGGCTTCCAAACCCCGTC 240
DB 277 TACCCGTCCTCCCGTCCGACGAGTCCCGGAGGTCCTGACGAGGCTTCCAAACCCCGTC 336

Db 277 TACCGTCCCCCGTGGCCCACTACGCCCCGAGGGTCTGACGACAGGCTTCCAAACCCGTC 336
Qy 241 GTCTGACGACAGCCCAATCCCATCC 267
Db 337 GTCTGACGACAGCCCAATCCCATCC 363

Search completed: January 31, 2003, 09:39:12
Job time : 76 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2003, 06:47:46 ; Search time 308 Seconds
(without alignments)
1952.219 Million cell updates/sec

Title: US-09-807-201-7

Sequence: 1 Gacattccagatattcattc.....cgcagcccaatccatcc 267

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq.10102.*

1:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
11:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
12:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
13:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
14:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
15:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
16:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
17:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
18:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
19:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
20:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
21:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA2001.DAT.*
22:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
23:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*
24:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA2004.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match length	ID	Description
1	267	100.0	267 21	AAZ95004
2	267	100.0	267 22	AAZ95004
3	267	100.0	267 22	AAZ95004
4	267	100.0	267 22	AAZ95004
5	267	100.0	267 22	AAZ95004
6	267	100.0	267 22	AAZ95004
7	267	100.0	267 22	AAZ95004
8	267	100.0	267 22	AAZ95004
9	267	100.0	267 22	AAZ95004

10	267	100.0	2479 22	AAZ95004	Human serine protease
11	267	100.0	2479 22	AAZ95004	Human transmembran
12	267	100.0	2479 24	AAZ95004	Prostate cancer-as
13	267	100.0	2479 24	AAZ95004	Human transmembran
14	267	100.0	2479 24	AAZ95004	Human transmembran
15	267	100.0	2479 24	AAZ95004	Human transmembran
16	267	100.0	2479 24	AAZ95004	Human transmembran
17	267	100.0	2479 24	AAZ95004	Human transmembran
18	267	100.0	2479 24	AAZ95004	Human transmembran
19	267	100.0	2479 24	AAZ95004	Human transmembran
20	267	100.0	2479 24	AAZ95004	Human transmembran
21	267	100.0	2479 24	AAZ95004	Human transmembran
22	267	100.0	2479 24	AAZ95004	Human transmembran
23	267	100.0	2479 24	AAZ95004	Human transmembran
24	267	100.0	2479 24	AAZ95004	Human transmembran
25	267	100.0	2479 24	AAZ95004	Human transmembran
26	267	100.0	2479 24	AAZ95004	Human transmembran
27	267	100.0	2479 24	AAZ95004	Human transmembran
28	267	100.0	2479 24	AAZ95004	Human transmembran
29	267	100.0	2479 24	AAZ95004	Human transmembran
30	267	100.0	2479 24	AAZ95004	Human transmembran
31	267	100.0	2479 24	AAZ95004	Human transmembran
32	267	100.0	2479 24	AAZ95004	Human transmembran
33	267	100.0	2479 24	AAZ95004	Human transmembran
34	267	100.0	2479 24	AAZ95004	Human transmembran
35	267	100.0	2479 24	AAZ95004	Human transmembran
36	267	100.0	2479 24	AAZ95004	Human transmembran
37	267	100.0	2479 24	AAZ95004	Human transmembran
38	267	100.0	2479 24	AAZ95004	Human transmembran
39	267	100.0	2479 24	AAZ95004	Human transmembran
40	267	100.0	2479 24	AAZ95004	Human transmembran
41	267	100.0	2479 24	AAZ95004	Human transmembran
42	267	100.0	2479 24	AAZ95004	Human transmembran
43	267	100.0	2479 24	AAZ95004	Human transmembran
44	267	100.0	2479 24	AAZ95004	Human transmembran
45	267	100.0	2479 24	AAZ95004	Human transmembran

ALIGNMENTS

RESULT 1	AAZ95004	standard; CDNA; 267 BP.
ID	AAZ95004	standard; CDNA; 267 BP.
AC	AAZ95004	
XX		
DT	15-AUG-2000	(first entry)
XX		
DE	CSG Prolis EST useful as prostate cancer marker.	
XX		
KW	Prostate cancer; cancer specific gene; CSG; expressed sequence tag;	
XX	EST; diagnosis; monitoring; staging; imaging; therapy; metastasis;	
KW	marker; human; Prolis; ss.	
XX		
OS	Homo sapiens.	
XX		
PN	MO200023111-A1.	
XX		
PD	27-APR-2000.	
XX		
PF	19-OCT-1999;	99MO-US24331.
XX		
PR	19-OCT-1999;	98US-0104737.
XX		
FA	(DIAD-) DIADEXUS LLC.	
XX		
PI	Salceda S, Recipon H, Catterkey R;	
XX		
DR	WPI; 2000-339531/29.	
XX		
PT	Diagnosing, staging and monitoring the presence and metastases of	
PT	prostate cancer especially useful for treating prostate cancer	

XX Human; prostate cancer; prostate-specific; diagnosis; vaccine;
KW cytosolic; gene therapy; metastasis; ss.

XX OS Homo sapiens.
 XX PN MO200151633-A2.
 XX PD 19-JUL-2001.
 XX PF 16-JAN-2001; 2001WO-US01574.
 XX PR 14-JAN-2000; 2000US-0483572.
 XX PA (CORI-) CORIXA CORP.
 XX PI Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG;
 XX PI Kalos MD, Fanger GR, Day CH, Retter MW, Stolk JA, Skeiky YAW,
 XX PI Wang A, Meagher MJ;
 XX PT WPI; 2001-425873/45.
 XX PS New polynucleotide encoding a prostate-specific protein, for
 XX PT diagnosing, monitoring and treating prostate cancer in a patient and
 XX PT for use in vaccines -
 XX PS Claim 1; Page 521; 543pp; English.
 XX CC The present invention describes polynucleotide sequences (I) which encode
 XX CC prostate-specific proteins (II) (I) and (II) have cytostatic activity,
 XX CC and can be used in vaccine production and gene therapy. (I), (II),
 XX CC antibodies to (II), fusion proteins comprising (II), and isolated
 XX CC T cells prepared using (I) or (II) are used to treat cancer in a patient.
 XX CC (I) and the antibodies are also used in the detection of cancer in a
 XX CC patient. The cancer that is diagnosed or treated is particularly
 XX CC prostate cancer. (I) and (II) can be used in vaccines. The antibodies or
 XX CC (I) can be used for monitoring the progression of cancer in a patient.
 XX CC (I) and (II) can also be used to improve diagnostic and therapeutic
 XX CC methods for prostate cancer. They can indicate the level of metastasis
 XX CC as well as the prostate volume. AAH93157 to AAH93944 and AA01115 to
 XX CC AA01318 represent polynucleotide and amino acid sequences used in the
 XX CC exemplification of the present invention.
 XX SO Sequence 683 BP; 160 A; 209 C; 171 G; 143 T; 0 other;
 Query Match 100.0%; Score 267; DB 22; Length 683;
 Best Local Similarity 100.0%; Pred. No. 3.3e-73;
 Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GAACATTCGATACCTATCATCTGATGCTGTGATTAACAGCAAGATGCTTGAAC 60
 DB 9 GAACATTCGATACCTATCATCTGATGCTGTGATTAACAGCAAGATGCTTGAAC 68
 QY 61 TCAAGGTCAACCAAGCTATTGACCTTACTATGAAAAACCATGATACCAACCGAAAAAC 120
 DB 69 TCAAGGTCAACCAAGCTATTGACCTTACTATGAAAAACCATGATACCAACCGAAAAAC 128
 QY 121 CCTATCCGACAGCGCCAGCTGATGCTGCTGCTAGAGGTGATCCGGCTCAGTAC 180
 DB 129 CCTATCCGACAGCGCCAGCTGATGCTGCTGCTAGAGGTGATCCGGCTCAGTAC 188
 QY 181 TACCGGTCCCGGTGCGCCAGTACGCGCCGAGAGGCTCTGACGCAAGGCTTCAACCCCGTC 240
 DB 189 TACCGGTCCCGGTGCGCCAGTACGCGCCGAGAGGCTCTGACGCAAGGCTTCAACCCCGTC 248
 QY 241 GTCTGACGACGACCCCAATCCCATCC 267
 DB 249 GTCTGACGACGACCCCAATCCCATCC 275
 RESULT 4
 ABL95536
 ID ABL95536 standard; cDNA; 683 BP.
 AC ABL95536;
 XX

DT 19-JUL-2002 (first entry)
 XX Human transmembrane protease serine 2 cDNA fragment SEQ ID NO 896.
 XX KM Human; cancer; prostate cancer; vaccine; cytostatic; immunostimulant;
 XX KM gene therapy; gene; ss.
 XX OS Homo sapiens.
 XX PN US2002022248-A1.
 XX PD 21-FEB-2002.
 XX PF 12-JAN-2001; 2001US-0759143.
 XX PR 25-FEB-1997; 97US-0806099.
 XX PR 01-AUG-1997; 97US-0904804.
 XX PR 09-FEB-1998; 98US-0020956.
 XX PR 25-FEB-1998; 98US-0030607.
 XX PR 14-JUL-1998; 98US-0115453.
 XX PR 23-SEP-1998; 98US-0159812.
 XX PR 15-JAN-1999; 99US-0232149.
 XX PR 09-APR-1999; 99US-0288946.
 XX PR 13-JUL-1999; 99US-0352616.
 XX PR 12-NOV-1999; 99US-0439313.
 XX PR 18-NOV-1999; 99US-0443686.
 XX PR 14-JAN-2000; 2000US-0483672.
 XX PR 27-MAR-2000; 2000US-0536857.
 XX PR 09-MAY-2000; 2000US-0568100.
 XX PR 12-MAY-2000; 2000US-0570737.
 XX PR 13-JUN-2000; 2000US-0593793.
 XX PR 27-JUN-2000; 2000US-0605783.
 XX PR 10-AUG-2000; 2000US-0636215.
 XX PR 29-AUG-2000; 2000US-0651236.
 XX PR 06-SEP-2000; 2000US-0657279.
 XX PR 02-OCT-2000; 2000US-0679426.
 XX PR 10-OCT-2000; 2000US-0685166.
 XX PA (XUJ/) XU J.
 XX PA (DILL/) DILLON D C.
 XX PA (MITC/) MITCHAM J L.
 XX PA (HARL/) HARLOCKER S L.
 XX PA (JIANG/) JIANG Y.
 XX PA (KALOS/) KALOS M D.
 XX PA (FANG/) FANGER G R.
 XX PA (RETT/) RETTER M W.
 XX PA (STOL/) STOLK J A.
 XX PA (DAYC/) DAY C H.
 XX PA (VEDV/) VEDVICK T S.
 XX PA (CART/) CARTER D.
 XX PA (LISK/) LI S X.
 XX PA (WANG/) WANG A.
 XX PA (SKEI/) SKEIKY Y A W.
 XX PA (HEPL/) HEPLER W T.
 XX PA (HEND/) HENDERSON R A.
 XX PI Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;
 XX PI Fanger GR, Retter MW, Stolk JA, Day CH, Vedvick TS, Carter D;
 XX PI Li SX, Wang A, Skeiky YAW, Hepler WT, Henderson RA;
 XX PT WPI; 2002-255649/30.
 XX PS New prostate-specific polynucleotides for diagnosing and treating
 XX PT diseases, in particular prostate cancer, and as markers for the
 XX PT progression of cancer -
 XX PS Claim 1; SEQ ID NO 896; 87pp; English.
 XX CC The present invention provides prostate-specific coding sequences and
 XX CC their encoded proteins. These can be used in the diagnosis and treatment
 XX CC of cancer, particularly prostate cancer. The present sequence is a cDNA
 XX CC described in the invention.

DR WPI, 2002-154967/20.
DR P-PSDB; AA818096.

XX Examining a biological sample for evidence of deregulated cellular
PT growth, comprises comparing the status of prostate-specific,
XX androgen-regulated, secreted serine protease, 20P1F12/TMPRSS2, in a
PT corresponding normal sample

XX Example 3; Fig 1; 161pp; English.

XX The present invention relates to methods and compositions for the
CC diagnosis and therapy of prostate, colon, bladder, lung, ovarian and
CC kidney cancer derived from or based on a normally prostate-specific,
CC androgen regulated, cell membrane associated secreted serine protease
CC termed 20P1F12/TMPRSS2. The invention further relates to a method of
CC examining a biological sample for evidence of deregulated cellular
CC growth comprising comparing the status of 20P1F12/TMPRSS2 gene (also
CC designated 20P1F12-GTCL) in the sample to the status of 20P1F12/TMPRSS2
CC in a corresponding normal sample. The invention also relates to 20P1F12/
CC TMPRSS2 polynucleotides and their corresponding proteins. Methods of the
CC invention are used for examining a sample such as blood, serum, stool,
CC urine, semen, or biopsy tissue for evidence of deregulated cell growth.
CC The deregulated cell growth is indicative of bladder cancer, lung
CC cancer, kidney cancer or ovarian cancer. It is useful for identifying
CC evidence of a neoplasm in a sample. Vaccines comprising an immunogenic
CC portion of 20P1F12/TMPRSS2 are useful for inhibiting growth of a cell
CC expressing 20P1F12/TMPRSS2 in a patient suffering from bladder cancer,
CC lung cancer, ovarian cancer or metastatic cancer. The present sequence
CC is human 20P1F12-GTCL cDNA.

XX Sequence 1738 BP; 433 A; 459 C; 473 G; 373 T; 0 other;

XX Query Match 100.0%; Score 267; DB 24; Length 1738;

XX Best Local Similarity 100.0%; Pred. No. 4.6e-73;

XX Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAACATTCAGATACCTATCATCTGATGCTGTGATGAACAGCAAGATGGCTTTGAC 60

DB 64 GAACATTCAGATACCTATCATCTGATGCTGTGATGAACAGCAAGATGGCTTTGAC 123

QY 61 TCAGGGTCAACACAGCTATTGACCTTATGATAAACCATGATACCAACCGGAAAC 120

DB 124 TCAGGGTCAACACAGCTATTGACCTTATGATAAACCATGATACCAACCGGAAAC 183

QY 121 CCTATCCCGACAGCCCACTGTGTCCTCCCACTGTCTACAGAGTGCATCCGGCTCAGTAC 180

DB 184 CCTATCCCGACAGCCCACTGTGTCCTCCCACTGTCTACAGAGTGCATCCGGCTCAGTAC 243

QY 181 TACCCGTCCCGTGCAGTACGCCCCGAGGGGTCTTACAGCAGAGGTTCCAAACCCCGTC 240

DB 244 TACCCGTCCCGTGCAGTACGCCCCGAGGGGTCTTACAGCAGAGGTTCCAAACCCCGTC 303

QY 241 GTCTGACGACAGCCCAATCCCATCC 267

DB 304 GTCTGACGACAGCCCAATCCCATCC 330

RESULT 7

ID AA290478 standard; cDNA; 2479 BP.

XX AA290478;

XX 06-JUN-2000 (first entry)

XX Ovr115 homolog protein encoding cDNA.

XX CSg; cancer specific gene; cancer; gynecologic cancer; ovarian; breast;

XX endometrial; uterine; lung; cytotoxic; ss.

XX Homo sapiens.

XX MO200012758-A1.

XX 09-MAR-2000.

XX 01-SEP-1999; 99MO-US19655.

XX 02-SEP-1998; 98US-0098880.

XX (DIAD-) DIADEXUS LLC.

XX Salceda S, Sun Y, Recipon H, Caffarely R;

XX WPI; 2000-256657/22.

XX P-PSDB; AA87280.

XX Diagnosing, staging, monitoring, imaging and treating cancer especially

XX PT gynecological cancers e.g. breast, ovarian cancer and lung cancer.

XX PT involves measuring cancer specific gene levels in cells and body fluids

XX Claim 9; Page 49-50; 58pp; English.

XX The invention relates to detecting, diagnosing metastasis and staging
CC cancer by measuring levels of cancer specific genes (CSG) in cells,
CC tissues or body fluids. Their remission and progression, decreases and
CC increases in CSG levels, is also monitored, by periodic sample analysis.
CC The methods are useful for detecting cancers, especially gynecologic
CC cancers which include ovarian, breast, endometrial and uterine cancer
CC and lung cancer. Antibodies against the CSgs labeled with paramagnetic
CC ions or a radioisotope is useful for imaging cancer and when conjugated
CC with a cytotoxic agent are useful for treating cancer. The present
CC sequence represents a Ovr115 homolog protein encoding cDNA, that
CC can be used for the detection of the various cancers.

XX Sequence 2479 BP; 578 A; 650 C; 677 G; 574 T; 0 other;

XX Query Match 100.0%; Score 267; DB 21; Length 2479;

XX Best Local Similarity 100.0%; Pred. No. 5.2e-73;

XX Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAACATTCAGATACCTATCATCTGATGCTGTGATGAACAGCAAGATGGCTTTGAC 60

DB 9 GAACATTCAGATACCTATCATCTGATGCTGTGATGAACAGCAAGATGGCTTTGAC 68

QY 61 TCAGGGTCAACACAGCTATTGACCTTATGATAAACCATGATACCAACCGGAAAC 120

DB 69 TCAGGGTCAACACAGCTATTGACCTTATGATAAACCATGATACCAACCGGAAAC 128

QY 121 CCTATCCCGACAGCCCACTGTGTCCTCCCACTGTCTACAGAGTGCATCCGGCTCAGTAC 180

DB 129 CCTATCCCGACAGCCCACTGTGTCCTCCCACTGTCTACAGAGTGCATCCGGCTCAGTAC 188

QY 181 TACCCGTCCCGTGCAGTACGCCCCGAGGGGTCTTACAGCAGAGGTTCCAAACCCCGTC 240

DB 189 TACCCGTCCCGTGCAGTACGCCCCGAGGGGTCTTACAGCAGAGGTTCCAAACCCCGTC 248

QY 241 GTCTGACGACAGCCCAATCCCATCC 267

DB 249 GTCTGACGACAGCCCAATCCCATCC 275

RESULT 8

ID AA287813 standard; DNA; 2479 BP.

XX AA287813;

XX 12-MAY-2000 (first entry)

XX Human tumour suppressor TMPRSS2 gene (GenBank Accn No: U75329).

XX Tumour suppressor gene; TMPRSS2; cancer; human; drug design;

XX gene therapy; protein therapy; ds.

PD		04-OCT-2001.	
XX			
PF	27-MAR-2001;	2001WO-US09919..	
PR	27-MAR-2000;	2000US-0536857.	
PR	09-MAY-2000;	2000US-0568100.	
PR	12-MAY-2000;	2000US-0570737.	
PR	13-JUN-2000;	2000US-0593793.	
PR	27-JUN-2000;	2000US-0605783.	
PR	10-AUG-2000;	2000US-0636215.	
PR	29-AUG-2000;	2000US-0651236.	
PR	06-SEP-2000;	2000US-0651279.	
PR	02-OCT-2000;	2000US-0679426.	
PR	10-OCT-2000;	2000US-0685166.	
XX			
PA	(CORI-) CORIXA CORP.		
XX			
P1	Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;		
P1	Fanger GR, Retter MW, Stoik JA, Day CH, Vedvick TS, Carter D;		
P1	Li SX, Wang A, Skelky YAW, Hepler WT, Henderson RA,		
XX			
DR	WPI; 2001-639232/73.		
XX	P-PStDB; AAU69939.		
PT	New human prostate-specific polypeptides and polynucleotides useful for		
PT	the diagnosis and treatment of cancer, especially prostate cancer -		
XX			
PS	Claim 1; Page 555-556; 579pp; English.		
XX			
CC	The invention relates to isolated prostate-specific		
CC	polynucleotides, polypeptides, fusion proteins of the polypeptides,		
CC	antibodies raised against the polypeptides (or antigenic epitopes		
CC	derived from them) and antigen-presenting cells expressing the		
CC	polypeptides. The antibodies are useful for detecting the presence of		
CC	cancer, especially prostate cancer. The polypeptides, polynucleotides and		
CC	the antigen-presenting cells are useful for stimulating and/or expanding		
CC	T cells specific for a tumour protein, and for inhibiting the development		
CC	of cancer especially prostate cancer. Compositions comprising the		
CC	polynucleotide and/or polypeptide are useful for stimulating an immune		
CC	response, and for treating cancer. The oligonucleotide is useful for		
CC	detecting cancer. The present sequence is a prostate specific		
CC	polynucleotide of the invention.		
XX			
SQ	Sequence 2479 BP; 578 A; 650 C; 677 G; 574 T; 0 other;		
	Query Match 100.0%; Score 267; DB 22; Length 2479;		
	Best Local Similarity 100.0%; Prod. No. 5.2e-77;		
	Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0		
OY	1 GAACATTCAGATACCTTACTGATGTGTTGAACAAGCATGGCTTTGAAC 60		
Db	9 GAACATTCAGATACCTTACTGATGTGTTGAACAAGCATGGCTTTGAAC 68		
OY	61 TCAGGTCACCAAGCATTGAGCCTTAACATAAGAAACATGATACCACCGAAAAC 120		
Db	69 TCAGGTCACCAAGCATTGAGCCTTAACATAAGAAACATGATACCACCGAAAAC 128		
OY	121 CCTATCCCAGACAGCCCACTGTGGTCCCACTGTCTACAGAGTGTCATCCGGCTCAGTAC 180		
Db	129 CCTATCCCAGACAGCCCACTGTGGTCCCACTGTCTACAGAGTGTCATCCGGCTCAGTAC 188		
OY	181 TACCGTCCCCCGTGCCTCAAGTAGGCCCGGAGGGTCTGACGACGAGCTTCCAACCCCGTC 240		
Db	189 TACCGTCCCCCGTGCCTCAAGTAGGCCCGGAGGGTCTGACGACGAGCTTCCAACCCCGTC 248		
OY	241 GTCTGACGACGAGCCCAATCCCATCC 267		
Db	249 GTCTGACGACGAGCCCAATCCCATCC 275		
RESULT 10			
AD	AADI3168 standard: DNA; 2479 BP.		
AD	AADI3168		

XX AAD13168;
 AC
 XX
 XX
 DT 16-OCT-2001 (first entry)
 XX
 DE Human serine protease DNA.
 XX
 KM Human; transmembrane serine protease; membrane-type serine protease;
 KM MTSP; protease domain; neoplastic disease; tumour; cancer; cytostatic;
 KM lung carcinoma; colon adenocarcinoma; ovarian carcinoma; gene therapy;
 KM serine protease; ds.
 XX
 XX Homo sapiens.
 OS
 FH Key Location/Qualifiers
 FT CDS 57..1535
 FT /tag= a
 FT /product= "Human serine protease"
 XX
 PN W0200157194-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 02-FEB-2001; 2001MO-US03471.
 XX
 PR 03-FEB-2000; 2000US-0179982.
 PR 18-FEB-2000; 2000US-0183542.
 PR 22-JUN-2000; 2000US-0213124.
 PR 26-JUL-2000; 2000US-0220970.
 PR 08-SEP-2000; 2000US-0657986.
 PR 22-SEP-2000; 2000US-0234840.
 XX
 XX (CORV-) CORVAS INT INC.
 XX
 PI Madison EL, Ong EO, Yeh J;
 DR P-PSDB; AAE06943.
 XX
 PT Novel single chain polypeptide comprising protease domain of type-II
 PT membrane-type serine protease or its catalytically active portion
 PT useful for treating and preventing cancer and tumor
 XX
 PS Disclosure; Page 250-252; 256pp; English.
 CC The invention relates to transmembrane serine proteases and their
 CC corresponding nucleotides and the protease domain of a type-II
 CC membrane-type serine protease (MTSP). MTSP is useful for identifying
 CC compounds that modulate or inhibit its proteolytic activity and for
 CC formulating a medicament for treating neoplastic disease. MTSP and
 CC its corresponding nucleotides are useful in preventing or treating
 CC tumours or cancers such as lung carcinoma, colon adenocarcinoma and
 CC ovarian carcinoma, in diagnostics and in hybridisation assays. MTSP
 CC is useful as a diagnostic marker for tumour development, growth and/or
 CC progression and as immunogens to generate antibodies that specifically
 CC bind to it. MTSP DNA is useful in a yeast two-hybrid system and in
 CC gene therapy. The present sequence is a DNA encoding human serine
 CC protease protein.
 CC
 SO Sequence 2479 BP; 578 A; 650 C; 677 G; 574 T; 0 other;
 XX
 QY Query Match 100.0%; Score 267; DB 22; Length 2479;
 DB Best Local Similarity 100.0%; Pred. No. 5,2e-73;
 Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GAACATTCGAGATACCTATCTACTGATGCTGTGATAACGACGATGGCTTGAAC 60
 DB 9 GAACATTCGAGATACCTATCTACTGATGCTGTGATAACGACGATGGCTTGAAC 68
 QY 61 TCAGGCTCACACACGATGATGACCTTACTATGAAAAACATGATCAACCGGAAAC 120
 DB 69 TCAGGCTCACACACGATGATGACCTTACTATGAAAAACATGATCAACCGGAAAC 128

QY 121 CCCTATCCCGACAGCCCACTGATGCTCCCACTGTCTACAGAGGATCCGGCTCACTAC 180
 DB 129 CCCTATCCCGACAGCCCACTGATGCTCCCACTGTCTACAGAGGATCCGGCTCACTAC 188
 QY 181 TACCCGTCCCGGAGCCCACTGATGCTCCCACTGTCTACAGAGGATCCGGCTCACTAC 240
 DB 189 TACCCGTCCCGGAGCCCACTGATGCTCCCACTGTCTACAGAGGATCCGGCTCACTAC 248
 QY 241 GTCTGACAGCAGCCCAATCCCATCC 267
 DB 249 GTCTGACAGCAGCCCAATCCCATCC 275
 XX
 XX RESULT 11
 ID AAH93928 standard; cDNA; 2479 BP.
 AC AAH93928;
 XX
 DT 04-OCT-2001 (first entry)
 XX
 DE Human transmembrane protease serine 2 full length cDNA sequence.
 XX
 KM Human; prostate cancer; prostate-specific; diagnosis; vaccine;
 KM cytostatic; gene therapy; metastasis; ss.
 OS Homo sapiens.
 XX
 PN W0200151633-A2.
 XX
 PD 19-JUL-2001.
 XX
 PF 16-JAN-2001; 2001MO-US01574.
 XX
 PR 14-JAN-2000; 2000US-0483672.
 XX
 PA (CORV-) CORVIA CORP.
 XX
 PI Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG;
 PI Kaloje MD, Fanger GR, Day CH, Retter MW, Stolk JA, Skelky YAW,
 PI Wang A, Meagher MJ;
 XX
 DR WPI; 2001-425873/45.
 XX
 PT New polynucleotide encoding a prostate-specific protein, for
 PT diagnosing, monitoring and treating prostate cancer in a patient and
 PT for use in vaccines
 XX
 PS Claim 1; Page 518-519; 543pp; English.
 CC The present invention describes polynucleotide sequences (I) which encode
 CC prostate-specific proteins (II), (I) and (II) have cytostatic activity,
 CC and can be used in vaccine production and gene therapy. (I), (II),
 CC antibodies to (II), fusion proteins comprising (II), and isolated
 CC T cells prepared using (I) or (II) are used to treat cancer in a patient.
 CC (I) and the antibodies are also used in the detection of cancer in a
 CC patient. The cancer that is diagnosed or treated is particularly
 CC prostate cancer. (I) and (II) can be used in vaccines. The antibodies or
 CC (I) can be used for monitoring the progression of cancer in a patient.
 CC (I) and (II) can also be used to improve diagnostic and therapeutic
 CC methods for prostate cancer. They can indicate the level of metastasis
 CC as well as the prostate volume. AAH93157 to AAH93944 and AAH01115 to
 CC AAH01318 represent polynucleotide and amino acid sequences used in the
 CC exemplification of the present invention.
 CC
 SO Sequence 2479 BP; 578 A; 650 C; 677 G; 574 T; 0 other;
 XX
 QY Query Match 100.0%; Score 267; DB 22; Length 2479;
 DB Best Local Similarity 100.0%; Pred. No. 5,2e-73;
 Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GAACATTCGAGATACCTATCTACTGATGCTGTGATAACGACGATGGCTTGAAC 60

Db 9 GAACATTCCAGATACCTATCTATCTGATGATACGCAAGATGGCTTGAAC 68
 Qy 61 TCAGGGTCACCAACCACTATTGACCTTACTATGAAAACCATGATACCAACCGGAAAAC 120
 Db 69 TCAGGGTCACCAACCACTATTGACCTTACTATGAAAACCATGATACCAACCGGAAAAC 128
 Qy 121 CCTATCCCGACAGCCCACTGTGTGCCCACTGTCTTACGAGGTGATCCGGCTCAGTAC 180
 Db 129 CCTATCCCGACAGCCCACTGTGTGCCCACTGTCTTACGAGGTGATCCGGCTCAGTAC 188
 Qy 181 TACCCGTCGCCCGGTGCGCCAGTACGCGCCGAGGGTCCCTACGAGGCTTCCAAACCCCGTC 240
 Db 189 TACCCGTCGCCCGGTGCGCCAGTACGCGCCGAGGGTCCCTACGAGGCTTCCAAACCCCGTC 248
 Qy 241 GTCTGCAGCAGCCCAATCCCATCC 267
 Db 249 GTCTGCAGCAGCCCAATCCCATCC 275

RESULT 12

ABK92201
 ID ABK92201 standard; DNA; 2479 BP.

AC ABK92201;

DT 15-AUG-2002 (first entry)

DE Prostate cancer-associated DNA sequence #87.

XX Prostate cancer; prostate tumour tissue; human; mammal; cytostatic;
 KM gene therapy; gene; ds.

OS Mammalia.

PN W0200230268-A2.

XX 18-APR-2002.

XX 12-OCT-2001; 2001WO-US32045.

XX 13-OCT-2000; 2000US-0687576.

XX 08-DEC-2000; 2000US-0733288.

XX 08-DEC-2000; 2000US-0733288.

XX 24-JAN-2001; 2001US-263957P.

XX 16-MAR-2001; 2001US-276791P.

XX 16-MAR-2001; 2001US-276888P.

XX 06-APR-2001; 2001US-281922P.

XX 24-APR-2001; 2001US-286214P.

XX 30-APR-2001; 2001US-0847046.

XX 04-MAY-2001; 2001US-288589P.

XX (BOSB-) EOS BIOTECHNOLOGY INC.

XX Gish KC, Mack DH, Wilson KE, Afar D, Hevez P;

XX WPI; 2002-471335/50.

XX P-PSDB; ABG61885.

XX Detecting a prostate cancer-associated transcript in a cell in a
 PT patient, useful for diagnosing prostate cancer (PC) or screening
 PT modulators of PC, by determining if prostate cancer-associated genes
 PT are expressed in a prostate tissue -

XX Claim 22; Page 371-372; 436pp; English.

XX The present invention relates to methods of detecting a prostate
 CC cancer-associated transcript in a cell from a patient. The method
 CC comprises contacting a biological sample from the patient with
 CC prostate cancer-associated polynucleotides (designated PC genes) that
 CC selectively hybridize to a sequence that is at least 80% identical
 CC to them. The prostate cancer-associated polynucleotide sequences
 CC are differentially expressed in prostate tumour tissue or in
 CC prostate cancer and are derived from the tissues of various

CC organisms such as humans or other mammals (e.g. mice, sheep and dogs).
 CC The methods of the invention are useful for diagnosing and treating
 CC prostate cancer in mammals. The prostate cancer-associated genes are
 CC useful for diagnosing or treating prostate cancer, as well as for
 CC identifying modulators of prostate cancer or agents that inhibit
 CC prostate cancer. The nucleic acid sequences are particularly useful
 CC in gene therapy, as a vaccine or in antisense applications.
 CC ABK92215-ABK92263 represent prostate cancer-associated polynucleotide
 CC sequences.

XX Sequence 2479 BP; 578 A; 650 C; 677 G; 574 T; 0 other.

XX Query Match 100.0%; Score 267; DB 24; Length 2479;

XX Best Local Similarity 100.0%; Pred. No. 5,2e-73;

XX Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAACATTCCAGATACCTATCTATCTGATGATACGCAAGATGGCTTGAAC 60
 Db 9 GAACATTCCAGATACCTATCTATCTGATGATACGCAAGATGGCTTGAAC 68
 Qy 61 TCAGGGTCACCAACCACTATTGACCTTACTATGAAAACCATGATACCAACCGGAAAAC 120
 Db 69 TCAGGGTCACCAACCACTATTGACCTTACTATGAAAACCATGATACCAACCGGAAAAC 128
 Qy 121 CCTATCCCGACAGCCCACTGTGTGCCCACTGTCTTACGAGGTGATCCGGCTCAGTAC 180
 Db 129 CCTATCCCGACAGCCCACTGTGTGCCCACTGTCTTACGAGGTGATCCGGCTCAGTAC 188
 Qy 181 TACCCGTCGCCCGGTGCGCCAGTACGCGCCGAGGGTCCCTACGAGGCTTCCAAACCCCGTC 240
 Db 189 TACCCGTCGCCCGGTGCGCCAGTACGCGCCGAGGGTCCCTACGAGGCTTCCAAACCCCGTC 248
 Qy 241 GTCTGCAGCAGCCCAATCCCATCC 267
 Db 249 GTCTGCAGCAGCCCAATCCCATCC 275

RESULT 13

ABL95535
 ID ABL95535 standard; cDNA; 2479 BP.

AC ABL95535;

DT 19-JUL-2002 (first entry)

XX Human transmembrane protease serine 2 cDNA sequence SEQ ID NO 894.

XX Human; cancer; prostate cancer; vaccine; cytostatic; immunostimulant;
 KM gene therapy; gene; ss.

XX Homo sapiens.

XX US2002022248-A1.

XX 21-FEB-2002.

XX 12-JAN-2001; 2001US-0759143.

XX 25-FEB-1997; 97US-0806099.

XX 01-AUG-1997; 97US-0904804.

XX 09-FEB-1998; 98US-0020956.

XX 25-FEB-1998; 98US-0030607.

XX 14-JUL-1998; 98US-0115453.

XX 23-SEP-1998; 98US-0159812.

XX 15-JAN-1999; 99US-0232149.

XX 09-APR-1999; 99US-0288946.

XX 13-JUL-1999; 99US-0352616.

XX 18-NOV-1999; 99US-0433913.

XX 14-JAN-2000; 2000US-0483672.

XX 27-MAR-2000; 2000US-0536857.

XX 09-MAY-2000; 2000US-0568100.

XX 12-MAY-2000; 2000US-0570737.

PR 13-JUN-2000; 2000US-0593793.
 PR 27-JUN-2000; 2000US-0605793.
 PR 10-AUG-2000; 2000US-0636215.
 PR 29-AUG-2000; 2000US-0651236.
 PR 06-SEP-2000; 2000US-0657279.
 PR 02-OCT-2000; 2000US-0679426.
 PR 10-OCT-2000; 2000US-0685166.
 XX
 PA (XUJ/) XU J.
 PA (DILL/) DILLON D. C.
 PA (MITC/) MITCHAM J. L.
 PA (HARL/) HARLOCKER S. L.
 PA (JIAN/) JIANG Y.
 PA (KALO/) KALOS M. D.
 PA (PANG/) PANGER G. R.
 PA (RETT/) RETTER M. W.
 PA (STOL/) STOLK J. A.
 PA (DAYC/) DAY C. H.
 PA (VEDV/) VEDVICK T. S.
 PA (CART/) CARTER D.
 PA (LISX/) LI S. X.
 PA (WANG/) WANG A.
 PA (SKET/) SKEIKY Y. A. W.
 PA (HEPL/) HEPLER W. T.
 PA (HEND/) HENDERSON R. A.
 PI Xu J., Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD,
 PI Fanger GR, Retter MW, Stolk JA, Day CH, Vedvick TS, Carter D;
 PI Li SX, Wang A, Skeiky YAW, Hepler WT, Henderson RA;
 DR WPI; 2002-255649/30.
 XX
 PT New prostate-specific polynucleotides for diagnosing and treating
 PT diseases, in particular prostate cancer, and as markers for the
 PT progression of cancer
 XX
 XX Claim 1; SEQ ID NO 894; 87bp; English.
 CC The present invention provides prostate-specific coding sequences and
 CC their encoded proteins. These can be used in the diagnosis and treatment
 CC of cancer, particularly prostate cancer. The present sequence is a CDNA
 CC described in the invention.
 XX
 XX Sequence 2479 BP; 578 A; 650 C; 677 G; 574 T; 0 other;
 SO
 Query Match 100.0%; Score 267; DB 24; Length 2479;
 Best Local Similarity 100.0%; Pred. No. 5.2e-73;
 Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GAACATTCAGATACCTATCATCTGATGCTGTGATTAACAGCAAGATGGCTTTGAAC 60
 DB 9 GAACATTCAGATACCTATCATCTGATGCTGTGATTAACAGCAAGATGGCTTTGAAC 68
 QY 61 TCAGGGTCAACCACTATCTGATGCTGTGATTAACAGCAAGATGGCTTTGAAC 120
 DB 69 TCAGGGTCAACCACTATCTGATGCTGTGATTAACAGCAAGATGGCTTTGAAC 128
 QY 121 CCTATCCCGCAGCAGCCACTGCTGCTCCCACTGTCTACAGAGGATCGGGCTCACTAC 180
 DB 129 CCTATCCCGCAGCAGCCACTGCTGCTCCCACTGTCTACAGAGGATCGGGCTCACTAC 188
 QY 181 TACCGTCCCGGTCGTCGCAAGTACGCGCCGAGGGGCTCTGACGAGGCTTCAACCCCGTC 240
 DB 189 TACCGTCCCGGTCGTCGCAAGTACGCGCCGAGGGGCTCTGACGAGGCTTCAACCCCGTC 248
 QY 241 GTCGTGACGAGCCCAAAATCCCATCC 267
 DB 249 GTCGTGACGAGCCCAAAATCCCATCC 275
 RESULT 14
 AAD28779
 ID AAD28779 standard; DNA; 2479 BP.

XX
 AC AAD28779;
 XX
 DT 07-MAY-2002 (first entry)
 XX
 DE Human TMPRSS2 gene.
 XX
 KW Serine protease; 20P1F12/TMPRSS2; 20P1F12-GTC1; cell growth; neoplasm;
 KW cancer; vaccine; human; TMPRSS2 gene; ds.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT 57..1535
 FT /tag= a
 FT /product= "Human TMPRSS2 protein"
 XX
 FN W0200204953-A2.
 XX
 PD 17-JAN-2002.
 XX
 PF 12-JUL-2001; 2001MO-US22168.
 XX
 PR 12-JUL-2000; 2000US-0615285.
 XX
 PA (AGEN-) AGENSYS INC.
 XX
 PI Saferan D, Rattano AB, Hubert RS, Jakobovits A, Paris M,
 PI Challita-Eld PM;
 XX
 DR WPI; 2002-154967/20.
 DR P-PSDB; AAE18097.
 XX
 PT Examining a biological sample for evidence of dysregulated cellular
 PT growth, comprises comparing the status of prostate-specific,
 PT androgen-regulated, secreted serine protease, 20P1F12/TMPRSS2, in a
 PT corresponding normal sample
 XX
 XX Example 3; Fig 2; 161bp; English.
 PS
 XX
 CC The present invention relates to methods and compositions for the
 CC diagnosis and therapy of prostate, colon, bladder, lung, ovarian and
 CC kidney cancer derived from or based on a normally prostate-specific,
 CC androgen regulated, cell membrane associated secreted serine protease
 CC termed 20P1F12/TMPRSS2. The invention further relates to a method of
 CC examining a biological sample for evidence of dysregulated cellular
 CC growth comprising comparing the status of 20P1F12/TMPRSS2 gene (also
 CC designated 20P1F12-GTC1) in the sample to the status of 20P1F12/TMPRSS2
 CC in a corresponding normal sample. The invention also relates to 20P1F12/
 CC TMPRSS2 polynucleotides and their corresponding proteins. Methods of the
 CC invention are used for examining a sample such as blood, serum, stool,
 CC urine, semen, or biopsy tissue for evidence of dysregulated cell growth.
 CC The dysregulated cell growth is indicative of bladder cancer, lung
 CC cancer, kidney cancer or ovarian cancer. It is useful for identifying
 CC evidence of a neoplasm in a sample. Vaccines comprising an immunogenic
 CC portion of 20P1F12/TMPRSS2 are useful for inhibiting growth of a cell
 CC expressing 20P1F12/TMPRSS2 in a patient suffering from bladder cancer,
 CC lung cancer, ovarian cancer or metastatic cancer. The present sequence
 CC is human TMPRSS2 gene.
 XX
 SO Sequence 2479 BP; 578 A; 650 C; 677 G; 574 T; 0 other;
 Query Match 100.0%; Score 267; DB 24; Length 2479;
 Best Local Similarity 100.0%; Pred. No. 5.2e-73;
 Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GAACATTCAGATACCTATCATCTGATGCTGTGATTAACAGCAAGATGGCTTTGAAC 60
 DB 9 GAACATTCAGATACCTATCATCTGATGCTGTGATTAACAGCAAGATGGCTTTGAAC 68
 QY 61 TCAGGGTCAACCACTATCTGATGCTGTGATTAACAGCAAGATGGCTTTGAAC 120
 DB 69 TCAGGGTCAACCACTATCTGATGCTGTGATTAACAGCAAGATGGCTTTGAAC 128

QY 121 CCTATCCCGCACAGCCCACTGTGTCCCACTGTCTTACGAGGTGATCCGGCTCAGTAC 180
 DB 129 CCTATCCCGCACAGCCCACTGTGTCCCACTGTCTTACGAGGTGATCCGGCTCAGTAC 188
 QY 181 TACCGGTCCCGGTGCGCCAGTACGCGCCGAGGGTCTTACGAGGCTTCCAAACCCCGTC 240
 DB 189 TACCGGTCCCGGTGCGCCAGTACGCGCCGAGGGTCTTACGAGGCTTCCAAACCCCGTC 248
 QY 241 GTCTGCACGACGCCCAAAATCCCATCC 267
 DB 249 GTCTGCACGACGCCCAAAATCCCATCC 275

RESULT 15

AAA08803
 ID AAA08803 standard; cDNA; 3245 BP.

AC AAA08803;

DT 01-AUG-2000 (first entry)

DE Androgen-inducible gene clone HrpCa6/7 found in LNCaP cells.

KW Androgen inducible; testosterone; prostate cancer; cytostatic;

KM TMPRSS2; diagnosis; ss.

OS Homo sapiens.

Key Location/Qualifiers

FT CDS 145..1623

FT /tag= a

FT /product= HrpCa6-7

PN MO200018961-A2.

PD 06-APR-2000.

PF 30-SEP-1999; 99WO-US22535.

PR 30-SEP-1998; 98US-0163759.

PR 30-SEP-1998; 98US-0164159.

PA (MILL-) MILLENNIUM PHARM INC.

PI Macbeth KJ, Shyjan AM,

DR WPI; 2000-293182/25.

DR P-PSDB; AAY92050.

PT Novel methods for identifying compounds for treating prostate cancer
 PT comprising measuring the level of expression or activity of 1 or more
 PT of 11 genes or their products

PS Claim 1; Fig 3; 108pp; English.

XX AAA08801-09 are genes which are androgen (e.g. testosterone) inducible
 CC in androgen-dependent prostate cancer cells (e.g. LNCaP cells) and
 CC constitutively expressed in androgen-independent prostate cancer cells
 CC (e.g. LNCaP cells). Agents which decrease the expression or
 CC activity of these clones may slow or arrest the growth of prostate cancer
 CC cells or may kill them. HrpCa6/7 can be obtained from the sequence of
 CC the known gene for TMPRSS2. A compound useful for treating prostate
 CC cancer can be identified in a novel method comprising measuring the
 CC expression level, or activity, of HrpCa2, 3, 6/7, 8, 9, 10, 13, 14, 15,
 CC 19, or peripheral-type benzodiazepine receptor (PBR) in a cell, in the
 CC presence and absence of a test compound. The sequences may also be used
 CC in diagnosis of prostate cancer and to determine efficacy of treatment
 CC for prostate cancer.

XX Sequence 3245 BP; 774 A; 818 C; 870 G; 783 T; 0 other;

XX Query Match

100.0%; Score 267; DB 21; Length 3245;

Best Local Similarity 100.0%; Pred. No. 5.8e-73;
 Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GAACATTCGAGTACTTATCATTTACTCGATGCTGTGATTAACGAGAATGCTTTGAAAC 60
 DB 97 GAACATTCGAGTACTTATCATTTACTCGATGCTGTGATTAACGAGAATGCTTTGAAAC 156
 QY 61 TCAGGGTACACCAACGACTATTGACTTACTATGAAAACCATGATATCCACCGGAAAAC 120
 DB 157 TCAGGGTACACCAACGACTATTGACTTACTATGAAAACCATGATATCCACCGGAAAAC 216
 QY 121 CCTATCCCGCACAGCCCACTGTGTCCCACTGTCTTACGAGGTGATCCGGCTCAGTAC 180
 DB 217 CCTATCCCGCACAGCCCACTGTGTCCCACTGTCTTACGAGGTGATCCGGCTCAGTAC 276
 QY 181 TACCGGTCCCGGTGCGCCAGTACGCGCCGAGGGTCTTACGAGGCTTCCAAACCCCGTC 240
 DB 277 TACCGGTCCCGGTGCGCCAGTACGCGCCGAGGGTCTTACGAGGCTTCCAAACCCCGTC 336
 QY 241 GTCTGCACGACGCCCAAAATCCCATCC 267
 DB 337 GTCTGCACGACGCCCAAAATCCCATCC 363

Search completed: January 31, 2003, 06:53:12
 Job time : 312 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2003, 06:47:46 ; Search time 2245 Seconds
(without alignments)
1926.145 Million cell updates/sec

Title: US-09-807-201-7

Sequence: 1 gaacattccagataccatc.....cgccagcccaatcccatcc 267

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapept 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST.*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estrov:*
6: em_estrpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_hiv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_oher:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match length	ID	Description
1	267	100.0	866	13 BI761763 603046751
2	209.8	78.6	300	9 AU099721 AU099721
3	127.6	47.8	514	10 AU044734 AU044734
4	127.2	47.6	345	9 AI613666 AI613666
5	127.2	47.6	398	9 AA451050 AA451050
6	127.2	47.6	456	9 AA832852 AA832852

7	127.2	47.6	489	10 AW912791 AW912791
8	127.2	47.6	555	10 AW211403 AW211403
9	127.2	47.6	583	10 BQ108875 BQ108875
10	127.2	47.6	723	13 B1850775 B1850775
11	127.2	47.6	783	12 BG176274 BG176274
12	127.2	47.6	838	13 B1144021 B1144021
13	127.2	47.6	915	13 B1143531 B1143531
14	127.2	47.6	944	14 BQ228388 BQ228388
15	127.2	47.6	1019	12 BE916909 BE916909
16	127.2	47.6	1137	13 B1687501 B1687501
17	125.6	47.0	756	10 AW211428 AW211428
18	125.6	47.0	796	13 B1556978 B1556978
19	125.6	47.0	921	13 B144727 B144727
20	118.8	44.5	741	13 B1683800 B1683800
21	117.8	44.1	846	12 BG870706 BG870706
22	116.8	43.7	486	10 AW909202 AW909202
23	112	41.9	790	12 BF144549 BF144549
24	110.2	41.3	338	12 BF001767 BF001767
25	105.4	39.5	497	12 BR857022 BR857022
26	92.2	34.5	633	12 BR816896 BR816896
27	86.8	32.5	934	12 BF788678 BF788678
28	82.6	30.9	177	12 BE696955 BE696955
29	79.4	29.7	523	12 BF190896 BF190896
30	77.2	28.9	599	13 B1105806 B1105806
31	69	25.8	782	12 BR675039 BR675039
32	69	25.8	1734	11 BC015819 BC015819
33	65.6	24.6	795	13 BG962521 BG962521
34	64	24.0	321	12 BF056592 BF056592
35	64	24.0	618	10 AW604314 AW604314
36	64	24.0	676	10 AW376911 AW376911
37	56	21.0	221	10 BB870649 BB870649
38	51.8	19.4	908	12 BE681173 BE681173
39	45.6	17.1	703	10 BE376601 BE376601
40	39.2	14.7	154	10 BB602684 BB602684
41	39.2	14.7	909	9 AF122350 AF122350
42	38.6	14.5	843	10 BE620434 BE620434
43	38	14.2	848	17 AG60231 AG60231
44	37.8	14.2	962	14 BQ715655 BQ715655
45	37.6	14.1	400	17 CNS03YR CNS03YR

ALIGNMENTS

RESULT 1
BI761763
LOCUS 866 bp mRNA linear EST 25-SEP-2001
DEFINITION 603046751F1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5187060 5',
BI761763
VERSION BI761763.1 GI:15753341
KEYWORDS
SOURCE EST.
ORGANISM human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 866)
REFERENCE
NIH-MGC http://mgs.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: L14M11467 row: e column: 13
High quality sequence stop: 781.
Location/Qualifiers
1. 866

FEATURES
Source

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5187060"
/clone_1ib="NIH_MGC_116"
/lab_host="DH10B"
/notes="Organ: pooled colon, kidney, stomach; Vector:
pcmv-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
source anonymous pool of 3 colons, age 26 yo male, 49 yo
female, 71 yo male colon; 46 yo male kidney, and pool of 2
stomachs, 62 yo male and 70 yo female. Library is
oligo-dt primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.4 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
023. Note: this is a NIH_MGC Library."

BASE COUNT      204 a      251 c      236 g      175 t
ORIGIN

Query Match      100.0%; Score 267; DB 13; Length 866;
Best Local Similarity 100.0%; Pred. No. 2.1e-63;
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAACATTCCAGATCCTATCTATCTGATGCTGTGTAACAAGCAAGATGGCTTGAAC 60
    |||
DB 70 GAACATTCAGATCCTATCTATCTGATGCTGTGTAACAAGCAAGATGGCTTGAAC 129
    |||

QY 61 TCAGGGTCACACAGTATTTGACCTTATGAAACCATGATACCAACCGGAAAC 120
    |||
DB 130 TCAGGGTCACACAGTATTTGACCTTATGAAACCATGATACCAACCGGAAAC 189
    |||

QY 121 CCTATCCGCAAGCCCACTGTGTGTCCTCTAGACAGGATGATCCGGCTCAGTAC 180
    |||
DB 190 CCTATCCGCAAGCCCACTGTGTGTCCTCTAGACAGGATGATCCGGCTCAGTAC 249
    |||

QY 181 TACCCGTCCTCCGTCGTCAGTACGCCCCAGAGGATCTCTGACGAGGCTTCCACCCCGTC 240
    |||
DB 250 TACCCGTCCTCCGTCGTCAGTACGCCCCAGAGGATCTCTGACGAGGCTTCCACCCCGTC 309
    |||

QY 241 GTCTGCACGACGCCCAATCCCATCC 267
    |||
DB 310 GTCTGCACGACGCCCAATCCCATCC 336
    |||

RESULT 2
A0099721      300 bp      mRNA      linear      EST 05-APR-2001
LOCUS
DEFINITION
A0099721 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
KAT01641 similar to Human serine protease mRNA, mRNA sequence.
ACCESSION
A0099721
VERSION
A0099721.1 GI:13550850
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 300)
Suzuki,Y., Tsunoda,T., Taiba,H., Mizushima-Sugano,J., Sese,J., Hata
,H., Oca,T., Isogai,T., Tanaka,T., Nakamura,Y., Morishita,S., Okubo
,K., Suyama,A. and Sugano,S.
In silico mapping of the 5'-ends of human mRNAs using full-length
enriched and 5'-end enriched cDNA libraries constructed by
Oligo-capping method
Unpublished (2001)
Contact: Yutaka Suzuki
Department of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minato-ku, Tokyo 108-8639, Japan
Email: ysuzuki@ms.u-tokyo.ac.jp
Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
,S. Construction and characterization of a full length-enriched and
a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
Location/Qualifiers

```

```

source      1. 300
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="KAT01641"
/clone_1ib="Sugano Homo sapiens cDNA library"
/notes="Differential display comparison of untreated and
dimethylfumarate treated U937 cells"

BASE COUNT      67 a      89 c      88 g      56 t
ORIGIN

Query Match      78.6%; Score 209.8; DB 9; Length 300;
Best Local Similarity 99.1%; Pred. No. 8.8e-48;
Matches 211; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAACATTCAGATCCTATCTATCTGATGCTGTGTAACAAGCAAGATGGCTTGAAC 60
    |||
DB 88 GAACATTCAGATCCTATCTATCTGATGCTGTGTAACAAGCAAGATGGCTTGAAC 147
    |||

QY 61 TCAGGGTCACACAGTATTTGACCTTATGAAACCATGATACCAACCGGAAAC 120
    |||
DB 148 TCAGGGTCACACAGTATTTGACCTTATGAAACCATGATACCAACCGGAAAC 207
    |||

QY 121 CCTATCCGCAAGCCCACTGTGTGTCCTCTAGACAGGATGATCCGGCTCAGTAC 180
    |||
DB 208 CCTATCCGCAAGCCCACTGTGTGTCCTCTAGACAGGATGATCCGGCTCAGTAC 267
    |||

QY 181 TACCCGTCCTCCGTCGTCAGTACGCCCCAGAG 213
    |||
DB 268 TACCCGTCCTCCGTCGTCAGTACGCCCCAGAG 300
    |||

RESULT 3
A0044734      514 bp      mRNA      linear      EST 18-SEP-1999
LOCUS
DEFINITION
A0044734 um13h07.y1 Sugano mouse kidney mRna Mus musculus cDNA clone
IMAGE:2192221 5' similar to SW:TM52_HUMAN O15393 TRANSMEMBRANE
PROTEASE, SERINE 2 ; , mRNA sequence.
ACCESSION
A0044734
VERSION
A0044734.1 GI:5905331
KEYWORDS
EST.
SOURCE
house mouse.
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 514)
Marras,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
Marras,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
Underwood,K., Stepien,M., Theising,B., Allen,M., Bowers,Y., Person
,B., Swaller,T., Gibbons,M., Page,D., Harvey,N., Schurr,R., Ritter
,E., Kohn,S., Shin,T., Jackson,T., Cardenas,M., McCann,R.,
Waterston,R. and Wilson,R.
The WashU-NCI Mouse EST Project 1999
Unpublished (1999)
Contact: Marra M/Mashu-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MG1:1004673
Seq primer: custom primer used
High quality sequence stop: 482.
Location/Qualifiers
1. 514
/organism="Mus musculus"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:2192221"
/clone_1ib="Sugano mouse kidney mRna"
/sex="female"
/dev stage="adult"
/lab_host="DH10B"

```

/note="Organ: Kidney; Vector: pME18S-FL3; Site 1: DraIII (CACTGTG); Site 2: DraIII (CACCATGTG); 1st strand cDNA was primed with an oligo(dT) primer (ATGGGCGCTTTTCTTTTCTTTT), double-stranded cDNA was ligated to a DraIII adaptor (TTGGGCGCTACTG), digested and cloned into distinct DraIII sites of the pME18S-FL3 vector (5' site CACTGTG, 3' site CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sunito Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTCTGTGCTTAAAGATGCG and 3' end primer CCACTGCGAGCTCGACACA."

BASE COUNT 125 a 139 c 147 g 103 t
ORIGIN

Query Match 47.8%; Score 127.6; DB 10; Length 514;
Best Local Similarity 71.0%; Pred. No. 6,7e-25;
Matches 169; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

27 CGATGCTGTGATACAGCAAGATGGCTTGAATCAGGGTCACACCAAGTATTGACC 86
95 CGGTGCTGCTGACACAGGACGATGGATGATCAAGGTCACCTCCAGGAATCGAAC 154
87 TTACTATGAAGAACATGATACCAACCGGAAACCCCTATCCGACAGCCCACTGTGT 146
155 TTGCTATGAGAACACAGGGATATAGTGTGACATCTGTCTCCGAGACACAGATGGC 214
147 CCCCATCTGTATAGAGGTGATCCGGCTCAGTACTACCCGCTCCCGCCAGTATGC 206
215 TCCCAATGGCTTAAACCTTGTATCCAGCCCACTACCTCCAGTGTCTCAGTATGC 274
207 CCCGAGGCTCTGACGAGGCTTCCAAACCCGCTGTGACGAGGCGCAATGCCA 264
275 TCCGAGATGACACAGCATGCTTCAACATCTGTATCCACACATCCCAAGTCTCA 332

RESULT 4
LOCUS A1613666 345 bp mRNA linear EST 21-APR-1999
DEFINITION v686g09.y1 Soares_mammary_gland_NBMWG Mus musculus cDNA clone
IMAGE:550720 5' similar to SW:TM52.HUMAN O15393 TRANSMEMBRANE
PROTEASE, SERINE 2; mRNA sequence.
A1613666
A1613666.1 GI:4622833
EST.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 345)
NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MG1:502872
This read is a RESEQUENCE of a previously sequenced mouse clone
This read has been verified (found to hit its original self in the
correct orientation)
Seq primer: -40RP from Gibco
High quality sequence stop: 280.
Location/Qualifiers
1. 345

FEATURES
source
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone_image="IMAGE:550720"
/clone_lib="Soares_mammary_gland_NBMWG"
/sex="male"

/tissue type="mammary gland"
/dev stage="4 weeks"
/lab_note="DH108"
/note="Organ: mammary gland; Vector: pT73D-Pac (Pharmacia)
) with a modified polylinker; Site 1: Not I; Site 2: Eco
RI; 1st strand cDNA was primed with a Not I - oligo(dT)
primer (5'
TCCTTACCATCTGCAAGTGGAGCGGCGGAGATGTTTCTTTTCTTTTCTTTT
T 3'); double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT73 vector.
RNA provided by Dr. Minoru Ko, Wayne State Univ. Library
constructed and normalized by Bento Soares and M. Fatima
Bonalda."

BASE COUNT 92 a 109 c 84 g 60 t
ORIGIN

Query Match 47.6%; Score 127.2; DB 9; Length 345;
Best Local Similarity 71.2%; Pred. No. 7.1e-25;
Matches 168; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

29 ATGCTGTGATTAACAGCAAGATGGCTTGAATCAGGGTCACACCAAGTATTGACCTT 88
82 ATGCTGCTGACACAGGACGATGGATGATGATGATGATGATGATGATGATGATGAT 141
89 ACTATGAAGAACATGATACCAACCGGAAACCCCTATCCGACAGCCCACTGTGTCC 148
142 GCTATGAGAACACAGGGATATAGTGTGACATCTGTCTCCGAGACACCAAGTGGCTC 201
149 CCACTGTCTGACAGGTCATCCGGCTCAGTACTACCCGCTCCCGGCGCCAGTATGCC 208
202 CCAATGCTTCAACCTTGTATCCAGCCCACTACCTCAATCTCCAGTGTCTCAGTATGC 261
209 CGAGGCTCTGACGAGGCTTCCAAACCCGCTGTGACGAGGCGCAATGCCA 264
262 CGAGATTAACAGCAAGCTCAACATCTGTATCCACACATCCCAAGTCTCA 317

RESULT 5
LOCUS AA451050 398 bp mRNA linear EST 04-JUN-1997
DEFINITION v686g09.r1 Soares_mammary_gland_NBMWG Mus musculus cDNA clone
IMAGE:550720 5' mRNA sequence.
AA451050
AA451050.1 GI:2164720
EST.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 398)
Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Giesel,S., Kucab,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
The WashU-HMI Mouse EST Project
Unpublished (1996)
Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MG1:502872
Seq primer: -28m3 rev2 ET from Amersham
High quality sequence stop: 397.
Location/Qualifiers
1. 398

FEATURES
source
/organism="Mus musculus"

```

/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:850720"
/clone_1ib="Soares_mammary_gland_NbMWG"
/sex="male"
/tissue_type="mammary gland"
/dev_stage="4 weeks"
/lab_host="DH10B"
/notes="Organ: mammary gland; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTACCAATCTGAAGTGGAGCGCGCGGAATGTTTTTTTTTTTTTTT T 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and Eco RI into the Not I and Eco RI sites of the modified pT73 vector. RNA provided by Dr. Minoru Ko, Wayne State Univ. Library constructed and normalized by Bento Soares and M. Fatima Bonaldo."
BASE COUNT      102 a      125 c      96 g      75 t
ORIGIN

```

```

Query Match      47.6%; Score 127.2; DB 9; Length 398;
Best Local Similarity 71.2%; Pred. No. 7.6e-25;
Matches 168; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 29 ATGCTGTGATTAACAGCAAGATGGCTTTGAATCAGGGTCACACAGCTATTGACCTT 88
DB 81 ATGCTGTGACACAGCAGAGATGATGATTAATCAGGTCATCTCCAGGAATCGACCTT 140
QY 89 ACTATGAAAACCATGATATCAACCGAAGAACCCCTATCCCGACAGCCCACTGTGTCC 148
DB 141 GCTATGAGAACCAACGGGTATCAGTCTGACACATCTGTCTCCGAGACACAGTGGCTC 200
QY 149 CCACTGTCTACAGAGTGCATCCGCTCACTACACCCGCTCCCGCCCACTAGACCCC 208
DB 201 CCAATGGCTCAACCTGTATTCAGCCCACTACCACTTCAGTCCCTCAGTATGCTC 260
QY 209 CGAGGCTCTGACAGAGCTTCCAAACCCCGTCTGACAGCAGCCCAATATCCCA 264
DB 261 CGAGATTAACAGCAGCCTCAACATCTGTATCCACACATCCCAAGTCTCA 316

```

```

RESULT 6
AA832852          456 bp      mRNA      linear      EST 23-FEB-1998
LOCUS             AA832852
DEFINITION       ub56g01.r1 Soares_mammary_gland_NMLMG Mus musculus cDNA clone
ACCESSION        AA832852
VERSION          AA832852.1 GI:2906580
KEYWORDS         EST.
SOURCE           house mouse.
ORGANISM         Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 456)
REFERENCE
AUTHORS         Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lemon, G., Soares, B., Wilson, R. and
Waterston, R.
TITLE           The WashU-HMI Mouse EST Project
JOURNAL          Unpublished (1996)
COMMENT          Contact: Maria M. Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LNL; contact the
IMAG Consortium (info@image.llnl.gov) for further information.
MG1:904436

```

```

Seq primer: -28m13 rev2 ET from Amerham
High quality sequence stop: 452.
Location/Qualifiers
FEATURES
SOURCE
1. 456
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:1381968"
/clone_1ib="Soares_mammary_gland_NMLMG"
/sex="female (lactating)"
/tissue_type="mammary gland"
/lab_host="DH10B"
/notes="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from mammary gland tissue from a lactating female, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and Eco RI into the Not I and Eco RI sites of the modified pT73 vector. Library is normalized. Library was constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT      99 a      138 c      114 g      105 t
ORIGIN

```

```

Query Match      47.6%; Score 127.2; DB 9; Length 456;
Best Local Similarity 71.2%; Pred. No. 8.1e-25;
Matches 168; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 29 ATGCTGTGATTAACAGCAAGATGGCTTTGAATCAGGGTCACACAGCTATTGACCTT 88
DB 27 ATGCTGTGACACAGCAGAGATGATGATTAATCAGGTCATCTCCAGGAATCGACCTT 86
QY 89 ACTATGAAAACCATGATATCAACCGAAGAACCCCTATCCCGACAGCCCACTGTGTCC 148
DB 87 GCTATGAGAACCAACGGGTATCAGTCTGACACATCTGTCTCCGAGACACAGTGGCTC 146
QY 149 CCACTGTCTACAGAGTGCATCCGCTCACTACACCCGCTCCCGCCCACTAGACCCC 208
DB 147 CCAATGGCTCAACCTGTATTCAGCCCACTACCACTTCAGTCCCTCAGTATGCTC 206
QY 209 CGAGGCTCTGACAGAGCTTCCAAACCCCGTCTGACAGCAGCCCAATATCCCA 264
DB 207 CGAGATTAACAGCAGCCTCAACATCTGTATCCACACATCCCAAGTCTCA 262

```

```

RESULT 7
AW912791          489 bp      mRNA      linear      EST 25-MAY-2000
LOCUS             AW912791
DEFINITION       uf45b02.y1 Soares_mammary_gland_NMLMG Mus musculus cDNA clone
ACCESSION        AW912791
VERSION          AW912791.1 GI:8078415
KEYWORDS         EST.
SOURCE           house mouse.
ORGANISM         Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 489)
REFERENCE
AUTHORS         NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE           National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL          Unpublished (1997)
COMMENT          Other ESTs: uf45b02.x1
Contact: Robert Strausberg, Ph.D.
Email: cgapdb-remail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAG Consortium (info@image.llnl.gov) for further information.
MG1:941135
Seq primer: -40RP from Glibco
High quality sequence stop: 443.
Location/Qualifiers
FEATURES
SOURCE
1. 489
/organism="Mus musculus"
/db_xref="taxon:10090"

```

/clone="IMAGE:1514283"
 /clone_id="Soares mammary_gland_NMLMG"
 /sex="female (lactating)"
 /tissue_type="mammary gland"
 /lab_host="MDH10B"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from mammary gland tissue from a lactating female, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT733 vector. Library is normalized. Library was constructed by Bento Soares and M. Patricia Bonaldo."

BASE COUNT 117 a 151 c 116 g 103 t 2 others
 ORIGIN

Query Match 47.6%; Score 127.2; DB 10; Length 489;
 Best Local Similarity 71.2%; Pred. No. 8.4e-25;
 Matches 166; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 29 ATGCTGTGATTAACAGCAAGATGGCTTGAACCTCAGGGTCACCAAGCTATTGACCTT 88
 DB 78 ATGCTGTGATTAACAGCAAGATGGCTTGAACCTCAGGGTCACCAAGCTATTGACCTT 137
 QY 89 ACTATGAAAACCATGATATCAACCGGAAAACCCCTATCCGCGACAGCCCACTGTGCTC 148
 DB 138 GCTATGAAACCAAGGATATCATGTGAGCAATCTGCTCCGAGACACCAAGTGGCTC 197
 QY 149 CCACTGTCTACAGATGATCCGGCTCAGTACTACCCGCTCCCGCCGACAGGCCC 208
 DB 198 CCAATGGCTCAACTTGTATCAAGCCCAAGTACTACCCATCTCCAGTCTCAGTATGCTC 257
 QY 209 CGAGGCTCTGACGACGAGCTTCAACCCCGCTGCTGACGAGCCCAAAATCCCA 264
 DB 258 CGAGGATTAACAGCAAGCTCAATCTGATTCACACACATCCCAAGTCTCA 313

RESULT 8
 LOCUS AW211403 555 bp mRNA linear EST 03-DEC-1999
 DEFINITION uc00e012.y1 NCI CGAP Mam3 Mus musculus cDNA clone IMAGE:2648878 5'
 similar to SW:FM52_HUMAN O15393 TRANSMEMBRANE PROTEASE, SERINE 2 ;
 mRNA sequence.
 ACCESSION AW211403
 VERSION AW211403.1 GI:6517358
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 555)
 REFERENCE NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
 AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 TITLE Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-remail.nih.gov
 Tissue Procurement: Lohar Hemmighausen Ph.D., Chu-Xia Deng Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CCAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 www-bio.llnl.gov/bdtp/image/image.html

MG1:1029330
 Seg primer: -40RP from Gibco
 High quality sequence stop: 427.
 Location/Qualifiers
 1..555
 /organism="Mus musculus"
 /strain="129,C57BL/6J,FVB/N"
 /db_xref="taxon:10090"

/clone="IMAGE:2648878"
 /clone_id="NCI CGAP Mam3"
 /tissue_type="tumor, gross tissue"
 /dev_stage="10 months"
 /lab_host="MDH10B"
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
 Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
 Library constructed by Life Technologies. Investigators
 providing samples: Lohar Hemmighausen/Chu-Xia Deng, NIH
 Reference for transgenic model: Xu et al., Nature Genetics
 22, 37-43 (1999)."

BASE COUNT 123 a 165 c 142 g 125 t
 ORIGIN

Query Match 47.6%; Score 127.2; DB 10; Length 555;
 Best Local Similarity 71.2%; Pred. No. 8.9e-25;
 Matches 168; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 29 ATGCTGTGATTAACAGCAAGATGGCTTGAACCTCAGGGTCACCAAGCTATTGACCTT 88
 DB 59 ATGCTGTGATTAACAGCAAGATGGCTTGAACCTCAGGGTCACCTCCAGGATCGACCTT 118
 QY 89 ACTATGAAAACCATGATATCAACCGGAAAACCCCTATCCGCGACAGCCCACTGTGCTC 148
 DB 119 GCTATGAAACCAAGGATATCATGTGAGCAATCTGCTCCGAGACACCAAGTGGCTC 178
 QY 149 CCACTGTCTACAGATGATCCGGCTCAGTACTACCCGCTCCCGCCGACAGGCCC 208
 DB 179 CCAATGGCTCAACTTGTATCAAGCCCAAGTACTACCCATCTCCAGTCTCAGTATGCTC 238
 QY 209 CGAGGCTCTGACGACGAGCTTCAACCCCGCTGCTGACGAGCCCAAAATCCCA 264
 DB 239 CGAGGATTAACAGCAAGCTCAATCTGATTCACACACATCCCAAGTCTCA 294

RESULT 9
 LOCUS B0108875 583 bp mRNA linear EST 16-APR-2002
 DEFINITION imageqc.8.2001/bmp14bdr81.y1 NCI CGAP Mam6 Mus musculus cDNA
 clone IMAGE:5342443 5', mRNA sequence.
 ACCESSION B0108875
 VERSION B0108875.1 GI:20158529
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 583)
 REFERENCE Kale,P.I., Harsch,T.J., Folta,P.A., Nelson,D.O., Sanders,C.G. and
 Prange,C.K.
 TITLE The I.M.A.G.E. Consortium quality control effort: clone
 JOURNAL resequencing for verification
 COMMENT Unpublished (2001)
 Other ESTs: B1683800
 Contact: Prange CK
 The I.M.A.G.E. Consortium
 Lawrence Livermore National Laboratory
 Livermore, CA, USA
 Email: help@image.llnl.gov
 This read has been verified (found to hit its original self in the
 correct orientation), as part of the I.M.A.G.E. Consortium quality
 control effort. High quality sequence is defined as having 100 or
 more base pairs with a phred quality value of 20 or greater, where
 or greater marks the beginning and end of the sequence. For
 information on obtaining this clone, please contact
 info@image.llnl.gov.
 Plate: LLML1869 row: o column: 20
 Seg primer: m3rpl
 High quality sequence stop: 583.
 Location/Qualifiers
 1..583
 /organism="Mus musculus"

FEATURES
 source
 1..583
 /organism="Mus musculus"

```

/strain="FVB/N"
/db_xref="taxon:10090"
/clone_image="5342443"
/clone_lib="NCI CGAP Mam6"
/sex="Female, virgin"
/issue_type="Infiltrating ductal carcinoma"
/dev_stage="5 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigator
providing samples: Jeffrey Green, M.D., NIH"
BASE COUNT      124 a 175 c 160 g 123 t
ORIGIN

```

```

Query Match      47.6%; Score 127.2; DB 14; Length 583;
Best Local Similarity 71.2%; Pred. No. 9.2e-25;
Matches 168; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 29 ATGCTGTTGATTAACAGAGATGGCTTTGAATCAGAGGTCCAGCAGCAGCTATTGACCTT 88
DB 213 ATGCTGTTGATTAACAGAGATGGCTTTGAATCAGAGGTCCAGCAGCAGCTATTGACCTT 272
QY 89 ACTATGAAAACCATGATACCAACCGGAAACCCCTATCCCGCAGACCCCACTGTGTCC 148
DB 273 GCTATGAGAACCAAGGGTATCACTGTGAGCACAATCTCTCCGAGACCAAGTGGCTC 332
QY 149 CCACTGTCTACAGAGGTGATCCGGCTCAGTACTACCCGTCGCCGCCAGTACGGCC 208
DB 333 CCAATGGCTTCAACTGTATCCAGCCAGTACTACCCATCTCCAGTGGCTCAGTATGCTC 392
QY 209 CGAGGGTCTCTGACGAGGCTTCCAAACCCGTCGTGTCAGCAGCAGCCCAATCCCA 264
DB 393 CGAGGATTAACAAGCAGCTCAACATCTGTCTATCCACACATCCCAAGTCTCTCA 448

```

```

RESULT 10      723 bp mRNA linear EST 10-OCT-2001
B1850775      imagegc.1.2001/smf23bdrxr81.y1 NCI CGAP Mam1 Mus musculus cdna
LOCUS          B1850775
DEFINITION     clone IMAGE:4461079 5', mRNA sequence.
ACCESSION      B1850775
VERSION        B1850775.1 GI:16004262
KEYWORDS       EST.
SOURCE         house mouse.
ORGANISM       Mus musculus

```

```

REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS        Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE          1 (bases 1 to 723)
JOURNAL        Kalle, P.I., Harsch, T.J., Folta, P.A., Nelson, D.O., Sanders, C.G. and
COMMENT        Prange, C.K.
                The I.M.A.G.E. Consortium quality control effort: clone
                resequencing for verification.
                Unpublished (2001)
                Other ESTs: BG176274
                Contact: Prange CK
                The I.M.A.G.E. Consortium
                Lawrence Livermore National Laboratory
                Livermore, CA, USA
                Email: help@image.llnl.gov
                This read has been verified (found to hit its original self in the
                correct orientation), as part of the I.M.A.G.E. Consortium quality
                control effort. High quality sequence is defined as having 100 or
                more base pairs with a phred quality value of 20 or greater, where
                a sliding window of 4 base pairs marks the beginning and end of the
                sequence. For information on obtaining this clone, please contact
                info@image.llnl.gov.
                Plate: L1AM10263 row: 1 column: 8
                Seq primer: m33p1
                High quality sequence stop: 723.
                Location/Qualifiers
                1..723
                /organism="Mus musculus"

```

```

/strain="FVB/N"
/db_xref="taxon:10090"
/clone_image="4461079"
/clone_lib="NCI CGAP Mam1"
/sex="Female, virgin"
/issue_type="tumor, biopsy sample"
/dev_stage="10 months, virgin"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"
BASE COUNT      177 a 197 c 190 g 155 t
ORIGIN

```

```

Query Match      47.6%; Score 127.2; DB 13; Length 723;
Best Local Similarity 71.2%; Pred. No. 1e-24;
Matches 168; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 29 ATGCTGTTGATTAACAGAGATGGCTTTGAATCAGAGGTCCAGCAGCAGCTATTGACCTT 88
DB 81 ATGCTGTTGATTAACAGAGATGGCTTTGAATCAGAGGTCCAGCAGCAGCTATTGACCTT 140
QY 89 ACTATGAAAACCATGATACCAACCGGAAACCCCTATCCCGCAGACCCCACTGTGTCC 148
DB 141 GCTATGAGAACCAAGGGTATCACTGTGAGCACAATCTCTCCGAGACCAAGTGGCTC 200
QY 149 CCACTGTCTACAGAGGTGATCCGGCTCAGTACTACCCGTCGCCGCCAGTACGGCC 208
DB 201 CCAATGGCTTCAACTGTATCCAGCCAGTACTACCCATCTCCAGTGGCTCAGTATGCTC 260
QY 209 CGAGGGTCTCTGACGAGGCTTCCAAACCCGTCGTGTCAGCAGCAGCCCAATCCCA 264
DB 261 CGAGGATTAACAAGCAGCTCAACATCTGTCTATCCACACATCCCAAGTCTCTCA 316

```

```

RESULT 11      783 bp mRNA linear EST 06-FEB-2001
BG176274      602338030P1 NCI CGAP Mam1 Mus musculus cdna IMAGE:4461079 5',
LOCUS          BG176274
DEFINITION     mRNA sequence.
ACCESSION      BG176274
VERSION        BG176274.1 GI:12682977
KEYWORDS       EST.
SOURCE         house mouse.
ORGANISM       Mus musculus

```

```

REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS        Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE          1 (bases 1 to 783)
JOURNAL        NIH-MGC http://mgc.nci.nih.gov/.
COMMENT        National Institutes of Health, Mammalian Gene Collection (MGC)
                Unpublished (1999)
                Contact: Robert Strausberg, Ph.D.
                Email: cgabbs-remail.nih.gov
                Tissue Procurement: Gilbert Smith, Ph.D.
                cDNA Library Preparation: Life Technologies, Inc.
                cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
                DNA Sequencing by: Incyte Genomics, Inc.
                Clone distribution: MGC clone distribution information can be
                found through the I.M.A.G.E. Consortium/LLNL at:
                http://image.llnl.gov
                Plate: L1AM10263 row: 1 column: 08
                High quality sequence stop: 665.
                Location/Qualifiers
                1..783
                /organism="Mus musculus"

```

```

/strain="FVB/N"
/db_xref="taxon:10090"
/clone_image="4461079"
/clone_lib="NCI CGAP Mam1"
/issue_type="tumor, biopsy sample"
/dev_stage="10 months, virgin"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;

```


Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH

BASE COUNT

188 a 213 c 211 g 171 t

Query Match

Best Local Similarity 47.6%; Score 127.2; DB 12; Length 783;
Matches 168; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

29 ATGCTGTGATTAACAGCAAGATGGCTTTGAATCTCAGGGTCACACCAAGCTATTGACCTT 88
Db ATGCTGTGACACAGCAGGATGATGATTAAGCTCAGGGTCACCTCCAGGAATCGGACCTT 140
QY ACTATGAAAACCATGATATCAACACCGAAGAACCCCTATCCGCGACAGCCCACTGGTCTC 148
Db GCTATGAGAACCAACGGGATATCATGCTGAGCACAATCTGCTCCGAGACCAACAGTGGCTC 200
QY 149 CCACTGTCTACAGAGGTGATCCGAGTCACTACCCGCTCCCGTCCGAGTACGCCC 208
Db 201 CCAATGGCTACAACTTGTATCAGCCAGTACTACCAATCTCCAGTGGCTCAGTATGCTC 260
QY 209 CGAGGCTCTACAGAGGCTTCCACCCGCTGCTGTCAGCAGGACCCCAATCCCA 264
Db 261 CGAGGATTACAACGAGCCTCAACATCTGTATCCACACATCCCAAGTCTCA 316

RESULT 12

LOCUS B1144021 838 bp mRNA linear EST 05-JUL-2001
DEFINITION 60290812AF1 NCI CGAP_Kid14 Mus musculus cDNA clone IMAGE:5065205
ACCESSION B1144021
VERSION B1144021.1 GI:14604022
KEYWORDS EST.

SOURCE

house mouse.
Mus musculus.

REFERENCE
AUTHORS Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Rodentia; Sciurognath; Muridae; Murinae; Mus.
JOURNAL NIH-MGC http://mgc.nci.nih.gov/
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LHAM1175 row: p column: 06
High quality sequence stop: 751.

FEATURES

source
Location/Qualifiers
1. 838
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:5065205"
/lab_host="NCI CGAP_Kid14"
/note="Organ: Kidney; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.75 kb. Constructed by Life
Technologies. Note: this is a NCI CGAP Library." | "

BASE COUNT

208 a 224 c 227 g 179 t

Query Match 47.6%; Score 127.2; DB 13; Length 838;
Best Local Similarity 71.2%; Pred. No. 1.1e-24;
Matches 168; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 29 ATGCTGTGATTAACAGCAAGATGGCTTTGAATCTCAGGGTCACACCAAGCTATTGACCTT 88
Db 87 ATGCTGTGACACAGCAGGATGATGATTAAGCTCAGGGTCACCTCCAGGAATCGGACCTT 146
QY 89 ACTATGAAAACCATGATATCAACACCGAAGAACCCCTATCCGCGACAGCCCACTGGTCTC 148
Db 147 GCTATGAGAACCAACGGGATATCATGCTGAGCACAATCTGCTCCGAGACCAACAGTGGCTC 206
QY 149 CCACTGTCTACAGAGGTGATCCGAGTCACTACCCGCTCCCGTCCGAGTACGCCC 208
Db 207 CCAATGGCTACAACTTGTATCAGCCAGTACTACCAATCTCCAGTGGCTCAGTATGCTC 266
QY 209 CGAGGCTCTACAGAGGCTTCCACCCGCTGCTGTCAGCAGGACCCCAATCCCA 264
Db 267 CGAGGATTACAACGAGCCTCAACATCTGTATCCACACATCCCAAGTCTCA 322

RESULT 13

LOCUS B1143513 915 bp mRNA linear EST 05-JUL-2001
DEFINITION 602907696F1 NCI CGAP_Kid14 Mus musculus cDNA clone IMAGE:5064815
ACCESSION B1143513
VERSION B1143513.1 GI:14603514
KEYWORDS EST.

SOURCE

house mouse.
Mus musculus.

REFERENCE
AUTHORS Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Rodentia; Sciurognath; Muridae; Murinae; Mus.
JOURNAL NIH-MGC http://mgc.nci.nih.gov/
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LHAM1174 row: o column: 24
High quality sequence start: 3
High quality sequence stop: 810.

FEATURES

source
Location/Qualifiers
1. 915
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:5064815"
/lab_host="NCI CGAP_Kid14"
/note="Organ: Kidney; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.75 kb. Constructed by Life
Technologies. Note: this is a NCI CGAP Library." | "

BASE COUNT

219 a 248 c 250 g 197 t

Query Match 47.6%; Score 127.2; DB 13; Length 915;
Best Local Similarity 71.2%; Pred. No. 1.1e-24;
Matches 168; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 29 ATGCTGTGATTAACAGCAAGATGGCTTTGAATCTCAGGGTCACACCAAGCTATTGACCTT 88
Db 162 ATGCTGTGACACAGCAGGATGATGATTAAGCTCAGGGTCACCTCCAGGAATCGGACCTT 221
QY 89 ACTATGAAAACCATGATATCAACACCGAAGAACCCCTATCCGCGACAGCCCACTGGTCTC 148
Db 222 GCTATGAGAACCAACGGGATATCATGCTGAGCACAATCTGCTCCGAGACCAACAGTGGCTC 281
QY 149 CCACTGTCTACAGAGGTGATCCGAGTCACTACCCGCTCCCGTCCGAGTACGCCC 208

Db 282 CCATGCTTACACTTATTCACGCCAGTACTCCATCTCCAGTCCCTCAGTATCTC 341
 QY 209 CGAGGGTCTGACGACGAGCTTCCAAACCCGTCGTCTGACGACGACCCAAATCCCA 264
 Db 342 CGAGGATTACAGCAGCAGCTCAACATCTGTCAATCCACACATCCCAAGTCTCA 397

RESULT 14
 BQ228388 944 bp mRNA linear EST 02-MAY-2002
 LOCUS BQ228388
 DEFINITION AGENCOURT 7578382 NCI CGAP Scl Mus musculus cDNA clone
 IMAGE:6051691 5', mRNA sequence.
 ACCESSION BQ228388
 VERSION BQ228388.1 GI:20409788
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 944)
 NIH-MGC http://mgs.nci.nih.gov/
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 plate: LLM1305 row: 0 column: 20
 High quality sequence stop: 707.

FEATURES
 source
 Location/Qualifiers
 1..944
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone_lib="IMAGE:6051691"
 /lab_host="NCI CGAP Scl"
 /note="Organ: stomach; Vector: pCMV-Sport6; Site 1: SalI;
 Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
 Average insert size 1.77 kb. Library constructed by Life
 Technologies. Note: this is a NCI CGAP Library."

BASE COUNT 231 a 243 c 263 g 207 t

Query Match 47.6%; Score 127.2; DB 14; Length 944;
 Best Local Similarity 71.2%; Pred. No. 1.2e-24;
 Matches 168; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 29 ATGCTGTATTAACAGCAAGATGGCTTGAATCAGGCTACCAACGATTTGACCTT 88
 Db 84 ATCTGTGTACACAGGAGATGGCATTTGAATCAAGGTACCTCCAGAAATGGACCTT 143
 QY 89 ACTATGAAACCATGATATCAACCGGAAACCCCTATCCCGACAGCCACTGTGCTC 148
 Db 144 GCTATGAGAACACGGGTATCACTGTGAGCAATCTCTCTCCGAGACCAACAGTGGCTC 203
 QY 149 CCACTGTCTACAGAGTGCATCCGGCTCAGTACTACCCGTCGCCCGGAGTACGCC 208
 Db 204 CCAATGGCTACAACTGTATTCAGCCAGCAATACCATCTCCAGTGCCTCAGTATGCTC 263
 QY 209 CGAGGGTCTGACGACGAGCTTCCAAACCCGTCGTCTGACGACGACCCAAATCCCA 264
 Db 264 CGAGGATTACAGCAGCAGCTCAACATCTGTCAATCCACACATCCCAAGTCTCA 319

RESULT 15
 BE916909 1019 bp mRNA linear EST 29-SEP-2000
 LOCUS BE916909

DEFINITION 60166450F1 NCI CGAP Mam1 Mus musculus cDNA clone IMAGE:3966301 5',
 mRNA sequence.
 BE916909
 ACCESSION BE916909
 VERSION BE916909.1 GI:10418022
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 1019)
 NIH-MGC http://mgs.nci.nih.gov/
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Gilbert Smith, Ph.D.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 plate: LLM9139 row: d column: 14
 High quality sequence stop: 600.

FEATURES
 source
 Location/Qualifiers
 1..1019
 /organism="Mus musculus"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone_lib="IMAGE:3966301"
 /clone_lib="NCI CGAP Mam1"
 /tissue_type="tumor, biopsy sample"
 /dev_stage="10 months, virgin"
 /lab_host="DH10B"
 /note="Organ: mammary; Vector: pCMV-Sport6; Site 1: SalI;
 Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
 Library constructed by Life Technologies. Investigator
 providing samples: Gilbert Smith, NIH"

BASE COUNT 281 a 318 c 261 g 159 t

Query Match 47.6%; Score 127.2; DB 12; Length 1019;
 Best Local Similarity 71.2%; Pred. No. 1.2e-24;
 Matches 168; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 29 ATGCTGTATTAACAGCAAGATGGCTTGAATCAGGCTACCAACGATTTGACCTT 88
 Db 59 ATGCTGTACACAGGAGATGGCATTTGAATCAAGGTACCTCCAGAAATGGACCTT 118
 QY 89 ACTATGAAACCATGATATCAACCGGAAACCCCTATCCCGACAGCCACTGTGCTC 148
 Db 119 GCTATGAGAACACGGGTATCACTGTGAGCAATCTCTCTCCGAGACCAACAGTGGCTC 178
 QY 149 CCACTGTCTACAGAGTGCATCCGGCTCAGTACTACCCGTCGCCCGGAGTACGCC 208
 Db 179 CCAATGGCTACAACTGTATTCAGCCAGCAATACCATCTCCAGTGCCTCAGTATGCTC 238
 QY 209 CGAGGGTCTGACGACGAGCTTCCAAACCCGTCGTCTGACGACGACCCAAATCCCA 264
 Db 239 CGAGGATTACAGCAGCAGCTCAACATCTGTCAATCCACACATCCCAAGTCTCA 294

Search completed: January 31, 2003, 08:28:28
 Job time : 2251 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: January 31, 2003, 06:47:46 ; Search time 3216 Seconds

(without alignments)
2416.184 Million cell updates/sec

Title: US-09-807-201-7

Perfect score: 1
Sequence: 1 gacattccagataccatc.....cgccagcccaatcccatcc 267

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

```

1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pac:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_scs:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: gb_vl:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pac:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_scs:*
28: em_un:*
29: em_vl:*
30: em_hcg_hum:*
31: em_hcg_inv:*
32: em_hcg_other:*
33: em_hcg_mus:*
34: em_hcg_pin:*
35: em_hcg_rnd:*
36: em_hcg_mam:*
37: em_hcg_vrt:*
38: em_sy:*
39: em_hgo_hum:*
40: em_hgo_mus:*
41: em_hgo_other:*

```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	* Query Match Length	DB ID	Description
1	267	100.0	683	AX201123
2	267	100.0	683	AX267922
3	267	100.0	1738	AX395316
4	267	100.0	1740	AF270487
5	267	100.0	2479	AX201121
6	267	100.0	2479	AX267965
7	267	100.0	2479	AX267920
8	267	100.0	2479	AX306771
9	267	100.0	2479	AX395318
10	267	100.0	2479	HSU75329
11	267	100.0	3226	AF329454
12	267	100.0	3245	AX201156
13	267	100.0	3245	AX267955
14	267	100.0	3966	AX041973
15	219	82.0	1476	AX267957
16	219	82.0	1476	AX267957
17	219	82.0	1479	AX201157
18	219	82.0	1479	AX267956
19	219	82.0	1479	AF133453
20	204	76.4	108927	AP001609
21	204	76.4	132110	AP001610
22	204	76.4	340000	HS21C085
23	178	66.7	196	AX106370
24	178	66.7	196	AX140651
25	178	66.7	196	AX200521
26	178	66.7	196	AX267177
27	127.2	47.6	1735	AF243500
28	127.2	47.6	1735	AF113596
29	127.2	47.6	1755	AF199362
30	118.4	44.3	1473	AB073550
31	103.4	38.7	220877	AC024957
32	64	24.0	2209	AK092554
33	48	18.0	111	AX155245
34	48	18.0	3660	HS2676404
35	38.8	14.5	158684	AC121032
36	38.6	14.5	10444	PVT422133
37	38.4	14.4	191325	AC125714
38	37.4	14.0	143798	AC012402
39	37.4	14.0	176802	AC078859
40	37.2	13.9	137390	AC113133
41	37.2	13.9	165725	AC024027
42	37.2	13.9	179676	AC015649
43	37	13.9	1641	RAB11A10
44	36.8	13.8	162336	AC021594
45	36.8	13.8	184510	AC079059

ALIGNMENTS

RESULT 1
AX201123
LOCUS AX201123 683 bp DNA linear PAT 29-AUG-2001
DEFINITION Sequence 753 from Patent WO0151633.
ACCESSION AX201123
VERSION AX201123.1 GI:15390897
KEYWORDS
SOURCE
ORGANISM human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 683)
REFERENCE Xu,J., Dillon,D.C., Micham,J.L., Harlocker,S.L., Jiang,Y.,
Reed,S.G., Kalos,M.D., Fanger,G.R., Day,C.H., Renter,M.W.,
Stolk,J.A., Shetty,Y.A., Wang,A. and Meagher,M.J.

RESULT 4	AF270487	LOCUS	DEFINITION	VERSION	ACCESSION	KEYWORDS
	AF270487	1740 bp	Homo sapiens androgen-regulated serine protease TMPRSS2 precursor (TMPRSS2) mRNA, complete cds.			
	AF270487					
	AF270487.1	GI:13540003				

SOURCE Homo sapiens.
ORGANISM Homo sapiens.
REFERENCE Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS 1 (bases 1 to 1740).
TITLE Afaar, D.E., Vivanco, I., Hubert, R.S., Kuo, J., Chen, E., Saffran, D.C., Raitano, A.B. and Jakobovits, A.
JOURNAL Catalytic cleavage of the androgen-regulated TMPRSS2 protease results in its secretion by prostate and prostate cancer epithelia
MEDLINE Cancer Res. 61 (4), 1686-1692 (2001)
PUBMED 21138112
11245484
2 (bases 1 to 1740)
REFERENCE Mitchell, S.C., Hubert, R.S. and Afaar, D.E.H.
AUTHORS Direct Submission
JOURNAL Submitted (19-MAY-2000) Urogenesis, Inc., 1701 Colorado Ave., Santa Monica, CA 90404, USA
FEATURES
source
Location/Qualifiers
1. .1740
/organism="Homo sapiens"
/db_xref="taxon:9606"
gene
1. .1740
/gene="TMPRSS2"
114. .1532
/gene="TMPRSS2"
/note="expressed in prostate and prostate cancer; differs from product encoded by the sequence presented in GenBank
Accession Number U75329"
/codon_start=1
/product="androgen-regulated serine protease TMPRSS2 precursor"
/protein_id="AAK29280.1"
/db_xref="GI:13540004"
/translation="MALNSGSPALGIPYEMNGYOPENPYPAQPTVTVVEVHPAY
YSPVQVAPRYLTQASNPVCTOPKSPGTVCTSKTKALCTLTGLFVLAALAA
GLMFMGSKNSGIECDSSGCIINPNCMDVSHCPGSDENRCVRLYGPFTIYV
YSSORSHAPVCOEDNNENYGRACMDYKNNFVSOCIIVDSGSTSMKNTASGN
VUIYKLYSDACSSKAVYSLACIACGNVLSRGRIVGBSALFGAMPQVSLHVO
NVHVGGSITTEWITTAHCEKELPNNMWHMTAFAGILRQSFMYAGYOVEKVI
ENYDRTKNDIALMKDQPLTFNDLVKVCCLPNNPQMLQPELCWISGMGATEEGK
TSEVNAKAVYLIEIQRNSRYVDLITPAMICAPFLQGVNDCGDSGGPLVTSKN
NTMTLIGDTSWCSGCAKAYRPVYGVWVFTDMYINQMBADG"
879. .1589
/gene="TMPRSS2"
/product="androgen-regulated serine protease TMPRSS2"
BASE COUNT 434 a 459 c 474 g 373 t
ORIGIN
Query Match 100.0%; Score 267; DB 9; Length 1740;
Best Local Similarity 100.0%; Pred. No. 6.2e-63;
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
mat_peptide
1 GAACATTCAGATACCTATCATCTGATGCTGTGATTAACGCAAGATGGCTTTGAAC 60
Db 66 GAACATTCAGATACCTATCATCTGATGCTGTGATTAACGCAAGATGGCTTTGAAC 125
Qy 61 TAGAGGTCAACCAAGCTATGAGCTTATGAAACCATGATGATCAACCGGAAAC 120
Db 126 TCAGGGTCAACCAAGCTATGAGCTTATGAAACCATGATGATCAACCGGAAAC 185
Qy 121 CCTATCCGCGACAGCCGCTGTCCTCCCACTGTCTAGAGGTGATCCGGCTGATAC 180
Db 186 CCTATCCGCGACAGCCGCTGTCCTCCCACTGTCTAGAGGTGATCCGGCTGATAC 245
Qy 181 TACCGGTCCCGGTGCGCCAGTACGCGGAGGCTCTGACGAGGCTTCAACCCCGTC 240
Db 246 TACCGGTCCCGGTGCGCCAGTACGCGGAGGCTCTGACGAGGCTTCAACCCCGTC 305
Qy 241 GTCTGACGACGACCAATCCCATCC 267
Db 306 GTCTGACGACGACCAATCCCATCC 332

RESULT 5
AX201121
LOCUS AX201121 2479 bp DNA linear PAT 29-AUG-2001
DEFINITION Sequence 751 from Patent WO0151633.
ACCESSION AX201121
VERSION AX201121.1 GI:15390895
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2479)
AUTHORS Xu, J., Dillon, D.C., Mitcham, J.L., Harlocker, S.L., Jiang, Y., Reed, S.G., Kalos, M.D., Fanger, G.R., Day, C.H., Retter, M.W., Stolk, J.A., Skeiky, Y.A., Wang, A. and Meagher, M.J.
TITLE Compositions and methods for the therapy and diagnosis of prostate cancer
JOURNAL Patent: WO 0151633-A 751 19-JUL-2001;
CORIXA CORPORATION (US)
FEATURES
source
Location/Qualifiers
1. .2479
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 578 a 650 c 677 g 574 t
ORIGIN
Query Match 100.0%; Score 267; DB 6; Length 2479;
Best Local Similarity 100.0%; Pred. No. 6.2e-63;
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GAACATTCAGATACCTATCATCTGATGCTGTGATTAACGCAAGATGGCTTTGAAC 60
Db 9 GAACATTCAGATACCTATCATCTGATGCTGTGATTAACGCAAGATGGCTTTGAAC 66
Qy 61 TCAGGGTCAACCAAGCTATGAGCTTATGAAACCATGATGATCAACCGGAAAC 120
Db 69 TCAGGGTCAACCAAGCTATGAGCTTATGAAACCATGATGATCAACCGGAAAC 128
Qy 121 CCTATCCGCGACAGCCGCTGTCCTCCCACTGTCTAGAGGTGATCCGGCTGATAC 180
Db 129 CCTATCCGCGACAGCCGCTGTCCTCCCACTGTCTAGAGGTGATCCGGCTGATAC 188
Qy 181 TACCGGTCCCGGTGCGCCAGTACGCGGAGGCTCTGACGAGGCTTCAACCCCGTC 240
Db 189 TACCGGTCCCGGTGCGCCAGTACGCGGAGGCTCTGACGAGGCTTCAACCCCGTC 248
Qy 241 GTCTGACGACGACCAATCCCATCC 267
Db 249 GTCTGACGACGACCAATCCCATCC 275
RESULT 6
AX207965
LOCUS AX207965 2479 bp DNA linear PAT 31-AUG-2001
DEFINITION Sequence 69 from Patent WO0157194.
ACCESSION AX207965
VERSION AX207965.1 GI:15422561
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2479)
AUTHORS Madison, E.L., Ong, E.O. and Yeh, J.C.
TITLE Nucleic acid molecules encoding transmembrane serine proteases, the encoded proteins and methods based thereon
JOURNAL Patent: WO 0157194-A 69 09-AUG-2001;
CORVUS INTERNATIONAL, INC. (US)
FEATURES
source
Location/Qualifiers
1. .2479
/organism="Homo sapiens"
/db_xref="taxon:9606"
CDS 57. .1535

```

/!note="Nucleotide sequence encoding human serine protease"
/codon_start=1
/protein_id="CAC60388.1"
/db_xref="GI:15422562"
/translation="MALNSGSPPAIGPYENHGYQENPYPAQTVVPEVHPAQY
VSPVPOVAPRVLTQASNPVCTQPKSPGCTCTSTKALCTITLGLFLVGAALAA
GLTKFMSKCSNGSIECDSSGTCINPSCGVSHCPGGBENRCVRLYGNFLLQ
YSSQKSWHPVQDDWNNENYGRACIDMKYKNNFYSSQGIYDSDSTSPMLKNTSAGN
VDIYKLYHSDSSKAVSLRCLACGVNLSSRSGIYVGSALPAGMPQVSLHVO
NVHVGCGSIITPEMIVTAHCEKPLNPMHTAFAGILRGSFMEYGAQYQVQYVSH
PNYDSKTRNDIALMKLOKPLTENDLVKVCILPNQPMLOPBLQCMISGWTERRKX
TSEVUNAAKVLITRQKNSRYVDNLITPMICAGFLQGNVDSQGDGGLVTSNN
NIMWLIGTSMGSCAKAYRPGVYGNVWFIDMIVYRKANG"
BASE COUNT      578 a      650 c      677 g      574 t
ORIGIN

```

```

Query Match      100.0%; Score 267; DB 6; Length 2479;
Best Local Similarity 100.0%; Pred. No. 6.2e-63;
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAACATTCAGATACCTATCTGATGCTGTGATGAACAGCAAGATGGCTTGAAC 60
DB 9 GAACATTCAGATACCTATCTGATGCTGTGATGAACAGCAAGATGGCTTGAAC 68
QY 61 TCAGGGTCAACACAGCTATGACCTTACTATGAAAACCATGATACCAACCGGAAAC 120
DB 69 TCAGGGTCAACACAGCTATGACCTTACTATGAAAACCATGATACCAACCGGAAAC 128
QY 121 CCTATCCGCAAGCCACTGTGCTCCCACTAGACAGGTGATCCGGCTCAGTAC 180
DB 129 CCTATCCGCAAGCCACTGTGCTCCCACTAGACAGGTGATCCGGCTCAGTAC 188
QY 181 TACCGTCCCCCGTCCGAGTACGCGCCGAGGGTCTTGAAGAGCTTCCACCCCGTC 240
DB 189 TACCGTCCCCCGTCCGAGTACGCGCCGAGGGTCTTGAAGAGCTTCCACCCCGTC 248
QY 241 GTCTGACGACGACCAATCCCATCC 267
DB 249 GTCTGACGACGACCAATCCCATCC 275

```

```

RESULT 7
LOCUS      AX267920      2479 bp      DNA      linear      PAT 26-OCT-2001
DEFINITION Sequence 894 from Patent WO0173032.
ACCESSION  AX267920
VERSION     AX267920.1 GI:16516526
KEYWORDS
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homimidae; Homo.
REFERENCE   1
AUTHORS    Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y.,
            Kalos,M.D., Fanger,G.R., Reiter,M.W., Stolk,J.A., Day,C.H.,
            Vedick,T.S., Carter,D., Li,S.X., Wang,A., Skelky,Y.A., Hepler,W.T.
            and Henderson,R.A.
TITLE      Compositions and methods for the therapy and diagnosis of prostate
            cancer
JOURNAL    Patent: WO 0173032-A 894 04-OCT-2001;
CORIXA CORPORATION (US)
FEATURES
SOURCE      1..2479
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
BASE COUNT  578 a      650 c      677 g      574 t
ORIGIN

```

```

Query Match      100.0%; Score 267; DB 6; Length 2479;
Best Local Similarity 100.0%; Pred. No. 6.2e-63;
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAACATTCAGATACCTATCTGATGCTGTGATGAACAGCAAGATGGCTTGAAC 60

```

```

DB 9 GAACATTCAGATACCTATCTGATGCTGTGATGAACAGCAAGATGGCTTGAAC 68
QY 61 TCAGGGTCAACACAGCTATGACCTTACTATGAAAACCATGATACCAACCGGAAAC 120
DB 69 TCAGGGTCAACACAGCTATGACCTTACTATGAAAACCATGATACCAACCGGAAAC 128
QY 121 CCTATCCGCAAGCCACTGTGCTCCCACTAGACAGGTGATCCGGCTCAGTAC 180
DB 129 CCTATCCGCAAGCCACTGTGCTCCCACTAGACAGGTGATCCGGCTCAGTAC 188
QY 181 TACCGTCCCCCGTCCGAGTACGCGCCGAGGGTCTTGAAGAGCTTCCACCCCGTC 240
DB 189 TACCGTCCCCCGTCCGAGTACGCGCCGAGGGTCTTGAAGAGCTTCCACCCCGTC 248
QY 241 GTCTGACGACGACCAATCCCATCC 267
DB 249 GTCTGACGACGACCAATCCCATCC 275

```

```

RESULT 8
LOCUS      AX306771      2479 bp      DNA      linear      PAT 11-DEC-2001
DEFINITION Sequence 4 from Patent WO0018961.
ACCESSION  AX306771
VERSION     AX306771.1 GI:17645937
KEYWORDS
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homimidae; Homo.
REFERENCE   1
AUTHORS    Macbeth,K.J. and Shyjan,A.W.
TITLE      Expression analysis of specific nucleic acids and polypeptides
            useful in the diagnosis and treatment of prostate cancer
JOURNAL    Patent: WO 0018961-A 4 06-APR-2000;
            Millennium Pharmaceuticals, Inc. (US)
FEATURES
SOURCE      1..2479
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /note="unnamed protein product"
            /codon_start=1
            /protein_id="CAD19059.1"
            /db_xref="GI:17645938"
CDS

```

```

/!note="Nucleotide sequence encoding human serine protease"
/codon_start=1
/protein_id="CAD19059.1"
/db_xref="GI:17645938"
/translation="MALNSGSPPAIGPYENHGYQENPYPAQTVVPEVHPAQY
VSPVPOVAPRVLTQASNPVCTQPKSPGCTCTSTKALCTITLGLFLVGAALAA
GLTKFMSKCSNGSIECDSSGTCINPSCGVSHCPGGBENRCVRLYGNFLLQ
YSSQKSWHPVQDDWNNENYGRACIDMKYKNNFYSSQGIYDSDSTSPMLKNTSAGN
VDIYKLYHSDSSKAVSLRCLACGVNLSSRSGIYVGSALPAGMPQVSLHVO
NVHVGCGSIITPEMIVTAHCEKPLNPMHTAFAGILRGSFMEYGAQYQVQYVSH
PNYDSKTRNDIALMKLOKPLTENDLVKVCILPNQPMLOPBLQCMISGWTERRKX
TSEVUNAAKVLITRQKNSRYVDNLITPMICAGFLQGNVDSQGDGGLVTSNN
NIMWLIGTSMGSCAKAYRPGVYGNVWFIDMIVYRKANG"
BASE COUNT      578 a      650 c      677 g      574 t
ORIGIN

```

```

Query Match      100.0%; Score 267; DB 6; Length 2479;
Best Local Similarity 100.0%; Pred. No. 6.2e-63;
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAACATTCAGATACCTATCTGATGCTGTGATGAACAGCAAGATGGCTTGAAC 60
DB 9 GAACATTCAGATACCTATCTGATGCTGTGATGAACAGCAAGATGGCTTGAAC 68
QY 61 TCAGGGTCAACACAGCTATGACCTTACTATGAAAACCATGATACCAACCGGAAAC 120
DB 69 TCAGGGTCAACACAGCTATGACCTTACTATGAAAACCATGATACCAACCGGAAAC 128
QY 121 CCTATCCGCAAGCCACTGTGCTCCCACTAGACAGGTGATCCGGCTCAGTAC 180
DB 129 CCTATCCGCAAGCCACTGTGCTCCCACTAGACAGGTGATCCGGCTCAGTAC 188

```

QY 181 TACCGGTCCCGGTCGAGTACGCGGCGGTCCTGAACGAGGCTTCCAAACCCCGTC 240
 Db 189 TACCGGTCCCGGTCGAGTACGCGGCGGTCCTGAACGAGGCTTCCAAACCCCGTC 248
 QY 241 GTCTGACGACGAGCCCAATCCCATCC 267
 Db 249 GTCTGACGACGAGCCCAATCCCATCC 275
 RESULT 9
 LOCUS AX395318 2479 bp DNA linear PAT 18-MAY-2002
 DEFINITION Sequence 3 from Patent WO0204953.
 ACCESSION AX395318
 VERSION AX395318.1 GI:21066314
 KEYWORDS human.
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1
 AUTHORS Saferan, D., Rattano, A.B., Hubert, R.S., Jakobovits, A., Farris, M. and
 Challita-Eid, P.M.
 TITLE Novel tumor antigen useful in diagnosis and therapy of bladder,
 ovary, lung and kidney cancers
 JOURNAL Patent: WO 0204953-A 3 17-JAN-2002;
 Agensys, Inc. (US)
 FEATURES
 source 1..2479
 location/Qualifiers
 /Organism="Homo sapiens"
 /db_xref="taxon:9606"
 BASE COUNT 578 a 650 c 677 g 574 t
 ORIGIN
 Query Match 100.0%; Score 267; DB 6; Length 2479;
 Best Local Similarity 100.0%; Pred. No. 6, 2e-63;
 Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GAACATCCGAGTACCTATCTAGTCTGATGATTAACAGCAAGTGGCTTTGAAC 60
 Db 9 GAACATCCGAGTACCTATCTAGTCTGATGATTAACAGCAAGTGGCTTTGAAC 68
 QY 61 TCAGGGTACACCAAGTATTTGACCTTATGAAAACCATGATACCAACCGGAAAC 120
 Db 69 TCAGGGTACACCAAGTATTTGACCTTATGAAAACCATGATACCAACCGGAAAC 128
 QY 121 CCTATCCGACAGCCCACTGTGTGCTTACGAGTGATCCGGCTCAGTAC 180
 Db 129 CCTATCCGACAGCCCACTGTGTGCTTACGAGTGATCCGGCTCAGTAC 188
 QY 181 TACCGGTCCCGGTCGAGTACGCGGCGGTCCTGAACGAGGCTTCCAAACCCCGTC 240
 Db 189 TACCGGTCCCGGTCGAGTACGCGGCGGTCCTGAACGAGGCTTCCAAACCCCGTC 248
 QY 241 GTCTGACGACGAGCCCAATCCCATCC 267
 Db 249 GTCTGACGACGAGCCCAATCCCATCC 275
 RESULT 10
 LOCUS HSU75329 2479 bp mRNA linear PRI 10-OCT-1997
 DEFINITION Human serine protease mRNA, complete cds.
 ACCESSION U75329
 VERSION U75329.1 GI:2507612
 KEYWORDS
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 2479)
 AUTHORS Paoloni-Giacobino, A., Chen, H., Peltsch, M.C., Rossier, C. and

TITLE Antomarakis, S.B.
 JOURNAL Cloning of the TMPRSS2 gene, which encodes a novel serine protease
 MEDLINE with transmembrane, LDIRA, and SRCR domains and maps to 21q22.3
 97468144
 PUBLISHED 9325052
 REFERENCE 2 (bases 1 to 2479)
 AUTHORS Paoloni-Giacobino, A., Chen, H. and Antomarakis, S.B.
 TITLE Direct Submission
 JOURNAL Submitted (17-OCT-1996) Medical Genetics, University of Geneva
 Medical School, 1 Michel-Servet Street, Geneva 1211, Switzerland
 FEATURES
 source 1..2479
 location/Qualifiers
 /Organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="21"
 /map="21q22.3"
 57..1535
 /codon_start=1
 /product="serine protease"
 /protein_id="AAC51784.1"
 /db_xref="GI:2507613"
 /translation="MALNSGSPPAIGPYENHGYOPENPYPAQPTVPTVEYHPAY
 YPSVPQYARVLTQASNPVCTOPKSPSGTCTSKTKALCTTLTGLTGLVGAALAA
 GLMKTGSKCSNSGIECDSSGTCINPSNCDGVSHCPGSEDNRCVRLKGFPILOM
 YSSORSKHPVCCDDWMENTGRACRDMGKNNFYSSQIVDSGTSFPMKLTSAQN
 VDIYKLYSHSDACSRAVSLRCLACGVNINSRSGRIYSGESALPGAMPQVSLHYO
 NHVCGGSIITPEMIYVAACHVEKPLNNPMTAPAGILRDSFMFYAGQYGVKXVSH
 PNYDSKTNNIDIALMKLQKPLFENDLYPVCLNPPGMLOPBLCMYISGMGABERGG
 TSEVLNAAKVLIETORCNSRYVDLITPAMICAFPLQGNVDSGDSGGLVTSNN
 NIMVLIGDTSWSSGCAKAPRGYGVNMFITDITIKOMKANG"
 BASE COUNT 578 a 650 c 677 g 574 t
 ORIGIN
 Query Match 100.0%; Score 267; DB 9; Length 2479;
 Best Local Similarity 100.0%; Pred. No. 6, 2e-63;
 Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GAACATCCGAGTACCTATCTAGTCTGATGATTAACAGCAAGTGGCTTTGAAC 60
 Db 9 GAACATCCGAGTACCTATCTAGTCTGATGATTAACAGCAAGTGGCTTTGAAC 68
 QY 61 TCAGGGTACACCAAGTATTTGACCTTATGAAAACCATGATACCAACCGGAAAC 120
 Db 69 TCAGGGTACACCAAGTATTTGACCTTATGAAAACCATGATACCAACCGGAAAC 128
 QY 121 CCTATCCGACAGCCCACTGTGTGCTTACGAGTGATCCGGCTCAGTAC 180
 Db 129 CCTATCCGACAGCCCACTGTGTGCTTACGAGTGATCCGGCTCAGTAC 188
 QY 181 TACCGGTCCCGGTCGAGTACGCGGCGGTCCTGAACGAGGCTTCCAAACCCCGTC 240
 Db 189 TACCGGTCCCGGTCGAGTACGCGGCGGTCCTGAACGAGGCTTCCAAACCCCGTC 248
 QY 241 GTCTGACGACGAGCCCAATCCCATCC 267
 Db 249 GTCTGACGACGAGCCCAATCCCATCC 275
 RESULT 11
 LOCUS AF329454 3226 bp mRNA linear PRI 16-MAY-2001
 DEFINITION Homo sapiens epithelialin (TMPRSS2) mRNA, complete cds.
 ACCESSION AF329454
 VERSION AF329454.1 GI:14091027
 KEYWORDS
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 3226)
 AUTHORS Jaguinet, E., Rao, N.V., Rao, G.V., Zhengming, W., Albertine, K.H. and
 Hoidal, J.R.

TITLE Cloning and characterization of the cDNA and gene for human epithelialin
JOURNAL Eur. J. Biochem. 268 (9), 2687-2699 (2001)
MEDLINE 2123025
PUBMED 11322890
REFERENCE 2 (bases 1 to 3226)
AUTHORS Jacquinet, E., Rao, N.V., Rao, G.V., Wang, Z., Albertine, K.H. and Hoidal, J.R.

TITLE Direct Submission
JOURNAL Submitted (15-DEC-2000) Pulmonary Medicine, University of Utah, 50N. Medical Dr., Salt Lake City, UT 84132, USA

FEATURES location/Qualifiers
source 1..3226
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="21"
 /map="21q22.3"
 1..3226
 /gene="TMPRSS2"
 128..1606
 /gene="TMPRSS2"
 /note="proteinase; PRSS10"
 /codon_start=1
 /product="epithelialin"
 /protein_id="AAK53559.1"
 /db_xref="GI:14091028"

gene
CDS
 128..1606
 /gene="TMPRSS2"
 /note="proteinase; PRSS10"
 /codon_start=1
 /product="epithelialin"
 /protein_id="AAK53559.1"
 /db_xref="GI:14091028"

BASE COUNT 782 a 808 c 857 g 779 t
ORIGIN
 1 GAACATTCAGATACCTATCATCTGATGCTGTGATGAACAGCAAGATGGCTTGAAC 60
 80 GAACATTCAGATACCTATCATCTGATGCTGTGATGAACAGCAAGATGGCTTGAAC 139
 61 TCAGGGTACCAACGCTATGACCTTACTATGAAACCAATGATGATGATGATGATGAT 120
 140 TCAGGGTACCAACGCTATGACCTTACTATGAAACCAATGATGATGATGATGATGAT 199
 121 CCTATCCCGACAGCCACTGTGCTCCCACTGTCTACAGAGTGCATCCGGCTCAGTAC 180
 200 CCTATCCCGACAGCCACTGTGCTCCCACTGTCTACAGAGTGCATCCGGCTCAGTAC 259
 181 TACCGGTCCCGGTCCCGGTCCCGGTCCCGGTCCCGGTCCCGGTCCCGGTCCCGGTCC 240
 260 TACCGGTCCCGGTCCCGGTCCCGGTCCCGGTCCCGGTCCCGGTCCCGGTCCCGGTCC 319
 241 GTCTGACGACGACCAATCCCATCC 267
 320 GTCTGACGACGACCAATCCCATCC 346

Query Match 100.0%; Score 267; DB 9; Length 3226;
Best Local Similarity 100.0%; Pred. No. 6.2e-63;
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 12
AX201156 3245 bp DNA linear PAT 29-AUG-2001
DEFINITION Sequence 786 from Patent WO0151633.
ACCESSION AX201156
VERSION AX201156.1 GI:15390911
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 (bases 1 to 3245)
AUTHORS Xu, J., Dillon, D.C., Mitcham, J.L., Harlocker, S.L., Jiang, Y., Reed, S.G., Kalos, M.D., Fanger, G.R., Day, C.H., Retter, M.W., Stolk, J.A., Skelky, Y.A., Wang, A. and Mesinger, M.J.
TITLE Compositions and methods for the therapy and diagnosis of prostate cancer
JOURNAL Patent: WO 0151633-A 786 19-JUL-2001;
FEATURES location/Qualifiers
source 1..3245
 /organism="Homo sapiens"
 /db_xref="taxon:9606"

BASE COUNT 774 a 818 c 870 g 783 t
ORIGIN

Query Match 100.0%; Score 267; DB 6; Length 3245;
Best Local Similarity 100.0%; Pred. No. 6.2e-63;
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAACATTCAGATACCTATCATCTGATGCTGTGATGAACAGCAAGATGGCTTGAAC 60
Db 97 GAACATTCAGATACCTATCATCTGATGCTGTGATGAACAGCAAGATGGCTTGAAC 156
Qy 61 TCAGGGTACCAACGCTATGACCTTACTATGAAACCAATGATGATGATGATGATGAT 120
Db 157 TCAGGGTACCAACGCTATGACCTTACTATGAAACCAATGATGATGATGATGATGAT 216
Qy 121 CCTATCCCGACAGCCACTGTGCTCCCACTGTCTACAGAGTGCATCCGGCTCAGTAC 180
Db 217 CCTATCCCGACAGCCACTGTGCTCCCACTGTCTACAGAGTGCATCCGGCTCAGTAC 276
Qy 181 TACCGGTCCCGGTCCCGGTCCCGGTCCCGGTCCCGGTCCCGGTCCCGGTCCCGGTCC 240
Db 277 TACCGGTCCCGGTCCCGGTCCCGGTCCCGGTCCCGGTCCCGGTCCCGGTCCCGGTCC 336
Qy 241 GTCTGACGACGACCAATCCCATCC 267
Db 337 GTCTGACGACGACCAATCCCATCC 363

RESULT 13
AX267955 3245 bp DNA linear PAT 26-OCT-2001
DEFINITION Sequence 929 from Patent WO0173032.
ACCESSION AX267955
VERSION AX267955.1 GI:16516540
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Buthera; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1
AUTHORS Xu, J., Dillon, D.C., Mitcham, J.L., Harlocker, S.L., Jiang, Y., Kalos, M.D., Fanger, G.R., Retter, M.W., Stolk, J.A., Day, C.H., Vedlick, T.S., Carter, D., Li, S.X., Wang, A., Skelky, Y.A., Hepler, W.T. and Henderson, R.A.

TITLE Compositions and methods for the therapy and diagnosis of prostate cancer

JOURNAL Patent: WO 0173032-A 929 04-OCT-2001;
FEATURES location/Qualifiers
source 1..3245
 /organism="Homo sapiens"
 /db_xref="taxon:9606"

BASE COUNT 774 a 818 c 870 g 783 t
ORIGIN

Query Match 100.0%; Score 267; DB 6; Length 3245;
Best Local Similarity 100.0%; Pred. No. 6.2e-63;
Matches 267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAACATTCAGATACCTATCATCTGATGCTGTGATGAACAGCAAGATGGCTTGAAC 60

Query Match	Best Local Similarity	100.0%	Score 2671	DB 6	Length 3966
Matches 267	Conservative 0	Mismatches 0	Indels 0	Gaps 0	
QY	1	GAACTTCCAGATACCACTTACTCGATCGTGTGATTAACGAGAAGTGGCTTTGAAC	60		
DB	9	GAACTTCCAGATACCACTTACTCGATCGTGTGATTAACGAGAAGTGGCTTTGAAC	68		
QY	61	TCAGGGTCACCAACAAGCTATTGACCTTACTATGAAAAACAATGATACCAACCGAAAAAC	120		
DB	69	TCAGGGTCACCAACAAGCTATTGACCTTACTATGAAAAACAATGATACCAACCGAAAAAC	128		
QY	121	CCCTATCCCGACAGGCCACCTGTGTCGCCACCTGTCTACAGAGTGCATTCGGCTCAGTAC	180		
DB	129	CCCTATCCCGACAGGCCACCTGTGTCGCCACCTGTCTACAGAGTGCATTCGGCTCAGTAC	188		

QY	49	ATGGGCTTTGAACCTCAGGGTCACACACAGAGCTATTGACCTTATCTATGAAGAACCATGATATAC	108
Dp	1	ATGGCTTTGAACCTCAGGGTCACACACAGAGCTATTGACCTTATCTATGAAGAACCATGATATAC	60
QY	109	CACCGGAAAAACCCCTTATCCCGACACGCCCATCTGTGATCCCATCTGTCTACGAGGTGCAT	168
Dp	61	CACCGGAAAAACCCCTTATCCCGACACGCCCATCTGTGATCCCATCTGTCTACGAGGTGCAT	120
QY	169	CCGGCTCAGTACTACCCGTCGCCCGTACGCCCAGTACGCCCGAGGGTCTTACGCAAGGCT	228
Dp	121	CCGGCTCAGTACTACCCGTCGCCCGTACGCCCAGTACGCCCGAGGGTCTTACGCAAGGCT	180
QY	229	TCCAAACCCCGTCGTCTGACAGCAGGCCCAATCCCATCC	287
Dp	181	TCCAAACCCCGTCGTCTGACAGCAGGCCCAATCCCATCC	219

Search completed: January 31, 2003, 07:50:50
Job time : 3222 secs